Naturalized formal epistemology of uncertain reasoning

Pfeifer, N.

Publication date:
2012

Naturalized formal epistemology
of uncertain reasoning

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op maandag 10 september 2012 om 10.15 uur door

NIKI PFEIFER,

geboren op 21 juni 1977 te Zams, Oostenrijk.

Promotiecommissie

Promotor: prof. dr. Stephan Hartmann
Copromotor: dr. Jan Sprenger

Overige leden: prof. dr. Igor Douven
prof. dr. Branden Fitelson
prof. dr. Reinhard Muskens
prof. dr. Gerhard Schurz

Financial support

This work is financially supported by the *Alexander von Humboldt Foundation*, the *German Research Foundation* project PF 740/2-1 "Rational reasoning with conditionals and probabilities. Logical foundations and empirical evaluation" (Project leader: Mag. Dr. Niki Pfeifer; Project within the DFG Priority Program SPP 1516 "New Frameworks of Rationality"), and the *Austrian Science Fund* project P20209 "Mental probability logic" (Project leader: Mag. Dr. Niki Pfeifer).

This work is dedicated

to my beloved wife

Katrin

Abstract

This thesis consists of a collection of five papers on *naturalized formal epistemology of uncertain reasoning*. In all papers I apply coherence based probability logic to make fundamental epistemological questions precise and propose new solutions to old problems. I investigate the rational evaluation of uncertain arguments, develop a new measure of argument strength, and explore the semantics of uncertain indicative conditionals. Specifically, I study formally and empirically the semantics of negated apparently self-contradictory conditionals (Aristotle's theses), resolve a number of paradoxes of the material conditional in a purely semantical way without employing pragmatics and investigate the psychological plausibility of the proposed semantics. Moreover, I defend the formalization of defeasible inferences within a probabilistic framework of nonmonotonic reasoning and empirically justify the formalizations by a series of psychological experiments. I investigate general properties of uncertain argument forms and the interrelations among logical validity, Adams' p-validity and probabilistic informativeness.

Contents

List of papers

As required by Article 17 of the PhD regulations of Tilburg University, the PhD candidate authored the papers or is the dominant co-author. The contributions of the PhD candidate to each paper are clearly indicated.

(p. 25) **Pfeifer, N.** (in press). On argument strength. In F. Zenker (Ed.), Workshop on Bayesian argumentation (special issue). Berlin, Heidelberg: *Synthese Library (Springer)*.
(Contribution of the PhD candidate: 100%.)

(p. 41) **Pfeifer, N.** (2012). Experiments on Aristotle's Thesis: Towards an experimental philosophy of conditionals. *The Monist, 95*(2), 223–240.
(Contribution of the PhD candidate: 100%.)

(p. 59) **Pfeifer, N.** & Kleiter, G. D. (2011). Uncertain deductive reasoning. In K. Manktelow, Over, D. E., and S. Elqayam (Eds.), *The science of reason: A Festschrift for Jonathan St B.T. Evans* (pp. 145-166). Hove, UK: Psychology Press.
(Contribution of the PhD candidate: 90%.)

(p. 76) **Pfeifer, N.** & Kleiter, G. D. (2010). The conditional in mental probability logic. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thought* (pp. 153-173). Oxford: Oxford University Press.
(Contribution of the PhD candidate: 85%.)

(p. 106) **Pfeifer, N.** & Kleiter, G. D. (2009). Framing human inference by coherence based probability logic. *Journal of Applied Logic, 7*(2), 206-217.
(Contribution of the PhD candidate: 65%.)

Chapter 1

Introduction

Naturalized epistemology was most prominently popularized by Quine (1969). It denotes a class of views that advocates empirical work to solve epistemological problems (Feldman, 2008). It is my firm belief that cognitive science and experimental psychology have a lot to say about central epistemological problems of reasoning under uncertainty. I will argue that formal work on central epistemological questions makes philosophical problems precise and allows for solutions that are at least hard to find by traditional (informal) epistemological methods. The use of conceptual analyses, which is common in traditional (informal) epistemology, brings the danger of "sometimes [employing] *exorbitantly speculative* examples or counterexamples" as philosophical arguments (Hendricks, 2006, p. ix; my emphasis). Psychological experiments on how people reason have a strong potential to empirically justify (counter)examples and thereby reduce substantially speculative elements of traditional conceptual analyses. I am therefore not claiming to replace traditional epistemology by formal epistemology, but rather advocate the use of formal tools wherever they fruitfully help to clarify the problem at hand.

The collection of papers in the present thesis attempts to provide new clarifications and solutions to a number of central epistemological questions. Specifically, in this thesis I

- elaborate a new measure of argument strength and rationally evaluate uncertain arguments w.r.t. the strength of arguments and the coherent probability propagation from the premises to the conclusion;
- investigate the semantics of uncertain indicative conditionals;
- analyze formally and empirically the semantics of negated apparently self-contradictory conditionals (Aristotle's theses);
- resolve a number of paradoxes of the material conditional in a purely semantical way without employing pragmatics and investigate the psychological plausibility of the proposed semantics;
- formalize defeasible inferences within a probabilistic framework of nonmonotonic reasoning and additionally justify empirically the formalizations; and
- investigate general properties of uncertain argument forms and the interrelations among logical validity, Adams' p-validity (1975, 1998) and probabilistic informativeness.

The normative elements described in the list above belong to the domain of formal epistemology. The descriptive or empirical elements are relevant for naturalized epistemologists. Moreover, the list points to experimental investigations on the underlying philosophical intuitions of the formal analyses. We recently argued that topics of the list above are candidates to extend the current domain of experimental philosophy (Pfeifer & Douven, submitted). Experimental philosophers investigated people's intuitions on a wide variety of philosophical topics, including causation, consciousness, cross-cultural intuitions, epistemology, folk morality, folk psychology, free will, and intentional action (see, e.g., Feltz, 2009; Knobe

& Nichols, 2008; Phillips, 2011). Uncertain reasoning, however, has been neglected so far. My thesis contributes to filling this gap.

There is a recent trend in formal epistemology (e.g., Bovens & Hartmann, 2003; Douven, 2008; Huber & Schmidt-Petri, 2009; Spohn, 2009) to use uncertainty measures for investigating conditionals and reasoning about uncertainty. There is a similar trend in the cognitive psychology and cognitive science of reasoning (e.g., Evans, in press; Oaksford & Chater, 2009; Over, Hadjichristidis, Evans, Handley, & Sloman, 2007; Pfeifer & Kleiter, 2009) to use probabilistic models for investigating human inference. Although both fields have different foci of interest I am convinced that the results and methodologies can be fruitfully applied in epistemology. The focus of the cognitive psychology of reasoning is on how people form representations of the premises and the conclusion and how they manipulate representations in order to draw inferences. The focus of the cognitive science of reasoning is on how to model human inferences. Both, how people reason and the computational modeling of human inference provide an empirical foundation for a naturalized epistemology of reasoning. My thesis elaborates a series of paradigmatic examples.

The present work consists of a collection of five full papers. One is accepted for publication and four are published. Among the various approaches to probability (Hájek, 2011), I have chosen *coherence based probability theory* as the common normative framework in each paper. The coherence approach to probability goes back to De Finetti (1980, 1974). More recent work includes, e.g., Walley (1991), Lad (1996), Biazzo and Gilio (2000), Coletti and Scozzafava (2002), and Galavotti (2008). The formal characterization of coherence is technically demanding. I therefore focus on the underlying philosophical intuitions in an informal way.[1] Intuitively, coher-

[1]Formally, coherence is characterized for example by Theorem 4 given by Coletti and Scozzafava (2002, p. 81), which is reproduced in the Appendix A (p. 132) below.

ence is in the tradition of subjective probability theory in which probabilities are conceived as *degrees of belief*. Degrees of belief are not only suitable for investigating epistemological problems but also for investigating empirical questions.

One key feature of coherence based probability theory is that the probability function is defined on an *arbitrary* family of conditional events. Therefore, a complete algebra is not required, which contrasts to the standard approach to probability. Conditional probability, $P(B \mid A)$, is a *primitive* notion. The probability value is assigned *directly* to the conditional event, $B \mid A$, as a whole (and not by definition via the fraction of the joint and the marginal probability, $P(A \wedge B)/P(A)$ if $P(A) > 0$). The conditional event $B \mid A$ is true if B is true and A is true, $B \mid A$ is false if B is false and A is true and $B \mid A$ is undetermined if A is false.

The probability axioms are formulated for conditional probabilities in the framework of coherence and not for absolute probabilities (as it is done in the standard approach to probability). Thus, "coherence" refers here to a foundation of probability theory. This meaning of "coherence" should not be confused with the epistemological problem of characterizing formally "how sentences hang together", which is also denoted by "coherence" (Bovens & Hartmann, 2003; Douven & Meijs, 2007).

Probability logical approaches became popular in formal epistemology (e. g., Adams, 1998; Haenni, Romeijn, Wheeler, & Williamson, 2011; Hailperin, 2000, 2011). Haenni et al. (2011, p. 3) point out that the fundamental problem of non-probabilistic logics consists in determining if a conclusion $\mathcal{C}$ is entailed by a premise set $\{\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n\}$. In contrast, probabilistic logics attach probabilities to the premises and the inference problem consists of determining what (set of) probabilities should be attached to the conclusion. For an example see Table 1.1. The books by Haenni et al. and by Hailperin focus on the fundamental problem of probability logic from

(1)	$P(A) = x$		A
(2)	$P(B \mid A) = y$		If A, then B
(3)	$xy \leq P(B) \leq xy + 1 - x$		B

Table 1.1: Probabilistic and non-probabilistic version of the *modus ponens*.

a normative point of view: what (set of) probabilities *should* be attached to the conclusion. This is also one of the central goals accomplished in my thesis. However, my thesis differs in two respects. Firstly, I use coherence based probability theory as a foundation of the proposed probability logic. Secondly, I study not only *normatively* the tightest coherent probability bounds[2] on the conclusion but also *descriptively* how people evaluate the probability bounds on the conclusion. The main goal of my work is to provide a new approach to uncertain reasoning that is normatively and descriptively adequate.

In my thesis, I propose coherence based probability theory as a normative framework for investigating probabilistic versions of propositional-logical argument forms (see Section 3.2–Section 3.5). *Coherence based probability logic* defines the consequence relation as a *deductive* one. The probabilistic inference problem consists of how to transmit the probabilities of the premises to the conclusion. Usually, the coherent probability of the conclusion is constrained by a lower and an upper probability.

As an example, consider the *modus ponens* argument form in Table 1.1. As only two probabilities are given in the premises (1) and (2), the coherent probability of the conclusion (3) is constrained by a lower and an upper probability bound.[3] Therefore, the conclusion of the probabilistic *modus*

[2]Coherent lower (l) and upper (u) probability bounds on conclusion $\mathcal{C}$ are *tight* if, and only if, they are the best possible coherent probability bounds on $\mathcal{C}$, i.e., there is no coherent probability less than l and greater than u.

[3]The law of total probability states that $P(B) = P(B|A) \times P(A) + P(B|\neg A) \times P(\neg A)$.

ponens is *imprecise*. If $P(B|\neg A) = z$ is added to the premise set of the probabilistic *modus ponens*, then the resulting argument form allows for inferring a *precise* probability value of the conclusion: $P(B) = xy + (1 - x)z$. The premises are not restricted to point probabilities but may also include imprecise probabilities and logical constraints. The former is important to formalize situations of incomplete probabilistic knowledge: The imprecise conclusion probability of a *modus ponens* inference may be used as a premise in a further argument. Moreover, if precise knowledge is not available to the reasoner during the initial probabilistic assessment of the premises, imprecise probabilities express a much more realistic state of knowledge compared with artificially forcing to guess precise values. Logical constraints may not only further constrain the tightest coherent probability bounds on the conclusion, they can also be used to express analytic truths, like "blue light is blue".

Coherence based probability logic (Pfeifer & Kleiter, 2006a, 2009) has received strong empirical support in a series of experiments on the rules of the nonmonotonic System P (Pfeifer & Kleiter, 2003, 2005, 2006b), indicative conditionals (Fugard, Pfeifer, Mayerhofer, & Kleiter, 2009, 2011; Fugard, Pfeifer, & Mayerhofer, 2011), and the conditional syllogisms (Pfeifer & Kleiter, 2007). My thesis substantially extends this work by formal and empirical investigations on Aristotle's Thesis (Section 3.2 p. 41), the paradoxes of the material conditional (Section 3.3 on p. 59), and on nonmonotonic argument forms (Section 3.4 on p. 76). Thus, my empirical work supports coherence based probability logic by external quality criteria which are beyond the purely formal ones (i.e., soundness and completeness).

All papers in this collection contribute to a naturalized formal epis-

As $P(\neg A) = 1 - P(A)$, the only unknown value is $P(B|\neg A)$. If $P(B|\neg A) = 0$, then $P(B) = P(B|A) \times P(A)$ (which is the tightest coherent lower probability bound of the conclusion). If $P(B|\neg A) = 1$, then $P(B) = P(B|A) \times P(A) + 1 - P(A)$ (which is the tightest coherent upper probability bound of the conclusion).

temology of uncertain reasoning. They are located in the intersection of formal epistemology, psychology and cognitive science of reasoning, and carefully extend the current domain of experimental philosophy to uncertain reasoning. The solid experimental work aims to avoid shallow empirical methodologies which—unfortunately—are still present in the recently emerged experimental philosophy movement. The next sections introduce the philosophical topics. Furthermore, they summarize and discuss selected main results in the light of their epistemological relevance.

Chapter 2

Summary of papers and outline of topics

2.1 Argument strength

An argument is an ordered triple consisting of (1) a (possibly empty) premise set, (2) a conclusion indicator[1], and (3) a conclusion. Classical logic tells us that if all premises are assumed to be true, then the conclusion is necessarily true if, and only if, the argument is logically valid. Logical validity guarantees truth preservation. Classical logic, however, does not tell us how to formalize uncertainty. The premises of everyday life arguments are usually uncertain and the conclusions are typically defeasible. Thus, a richer formal theory is needed for the rational evaluation of everyday life arguments.

In Section 3.1 (p. 25) I explore a new measure of argument strength that serves to rationally evaluate uncertain everyday life arguments. Let x denote the tightest coherent lower and y denote the tightest coherent upper bound on the conclusion of an argument $\mathcal{A}$. Then the argument strength of

[1] "Therefore", "hence", and "it follows that" are examples of conclusion indicators frequently used in commonsense arguments.

$\mathcal{A}$ is equal to

$$[1-(y-x)] \times \frac{x+y}{2}\,.$$

The measure maps onto the real interval of zero and one, where zero is the lowest value and one is the highest value of the argument strength.

The strength of $\mathcal{A}$ increases if the precision of the probability is high and the tightest coherent probability interval approaches 1. Thus, instead of requiring the assumption of unconditional truth of the premises, the premises are evaluated probabilistically and the tightest coherent probability bounds on the conclusion are used to assess the rationality of the argument.

This measure contrasts with traditional measures of confirmation (see, e. g. Crupi, Tentori, & Gonzales, 2007) and intuitions on argument strength discussed in the cognitive science literature (e. g., Rips, 2001; Hahn & Oaksford, 2006). In my proposal, the consequence relation remains deductive, while alternative proposals assume an uncertain relation between the premises and the conclusion. The advantage of my measure is that it does justice to the logical structure of the premise set: it is easy to represent conditionals. If a measure of argument strength requires to calculate the conditional probability of the conclusion given some combination of the premises, $P(\text{conclusion} \,|\, \text{premise set})$, then the severe problem arises of how to combine non-truth functional conditionals. While it is straightforward to formulate conjunctions of material conditionals or embedded material conditionals, it is not clear how to combine or conditionalize on conditional events (for a proposal see, e.g., Douven, 2012). It is hard to impose a satisfactory semantics on expressions like "$B \mid [A \wedge (B|A)]$" (i. e., a form of *modus ponens*, consisting of the conclusion B and the premise set $\{A, B|A\}$). In my proposal this problem is avoided, as probability logic tells us how to infer the tightest coherent probability bounds of the conclusion from the premises, which are in turn exploited for calculating the argument strength.

The fundamental problem of conditionalizing on conditionals is often ignored but virulent in current debates of establishing measures of "coherence among sentences" in the sense of "sentences that hang together" (for an overview on coherence measures see Douven & Meijs, 2007). Future work is needed to explore if coherence based probability logic provides the appropriate tools to come up with a new and satisfactory measure of coherence among sentences.

2.2 Aristotle's theses

One version of Aristotle's Thesis consists of the following negated conditional:

It is not the case that: If A, then not-A.

The other version of Aristotle's Thesis is obtained by replacing the conditional by "If not-A, then A".

Both versions are logically contingent if the conditional is interpreted as a material one, $A \supset \neg A$ and $\neg A \supset A$, respectively. Therefore, if the reasoner is completely uncertain about A, then both formulæ are evaluated by probabilistically non-informative intervals, $P(\neg(A \supset \neg A)) \in [0,1]$ and $P(\neg(\neg A \supset A)) \in [0,1]$, respectively. However, if the conditionals are interpreted as conditional events—$\neg A|A$ and $A|\neg A$, respectively—then coherence requires that the corresponding formulæ should necessarily obtain probability one. The material conditional interpretation seems counterintuitive, whereas the conditional event interpretation matches with commonsense intuitions.

This probability logical analysis is elaborated in Section 3.2 (p. 41). Furthermore, this section reports for the first time experiments on both forms of Aristotle's theses. I rationally reconstruct how people interpret and negate indicative conditionals. One main result consists in the empirical observation that people interpret indicative conditionals as conditional events. Another main result is that people negate conditionals by the narrow scope reading of the negation of conditionals (i.e., by negating the consequent of the conditional). Neither wide ($\neg(A \supset C)$) nor narrow scope ($(A \supset \neg C)$) readings of the negated material conditional are empirically supported. I argue for extending the domain of experimental philosophy to uncertain conditionals in Section 3.2.

In Pfeifer (2011b) I make a case for normative models by pointing out that they are necessary for the study of reasoning and furthermore, that systematic rationality norms provide research roadmaps and clarity. Formal work makes psychological theories precise and stimulates new psychological hypotheses. The fine-graded differences between the narrow and wide scope readings and the semantical implications of negating conditionals are hard to see without a formal background theory. Empirical work provides extra quality criteria of formal theories and thereby helps to arbitrate between formal theories beyond purely formal quality criteria like soundness and completeness. Section 3.2 and the subsequent sections make a strong case for the fruitful interactions of formal and empirical work.

Section 3.2 shows that Aristotle's theses provide formal and empirical litmus tests for contrasting the material conditional with the conditional event interpretation of indicative conditionals. Moreover, Section 3.2 sheds new light on how conditionals should be negated and on how people negate conditionals.

2.3 Paradoxes of the material conditional

The conditional event interpretation of indicative conditionals matches in numerous scenarios with commonsense intuitions, whereas the material conditional interpretation fails. Aristotle's theses is one class of examples (see the previous Section 2.2). Another class of examples consists of the so-called "paradoxes of the material conditionals". The paradoxes arise if natural language versions of indicative conditionals are formalized by the material conditional. "(Bill Gates is bankrupt) $\supset$ (Bill Gates is a billionaire)" follows logically from "Bill Gates is a billionaire", which is intuitively implausible.

Section 3.3 (p. 59) investigates conditional introduction inferences from affirmed consequents (Paradox 1)[2] and negated antecedents (Paradox 2) in a probabilistic setting. If conditionals are interpreted as material conditionals, then high premise probabilities guarantee high conclusion probabilities in both paradoxes. If, however, conditionals are interpreted as conditional events, then the unit interval is a coherent probability assessment of the conclusion. If it is the case that logical or probabilistic relations between the involved sentences are known (e. g., that bankruptcy logically excludes being a billionaire), then such information needs to be made explicit by accordingly augmenting the premise set. Such strengthening of the premises can lead to probabilistically informative arguments (Pfeifer & Douven, submitted). If, however, the logical and probabilistic relations between the involved sentences are unknown, then coherence requires that the probability of the conclusion of the paradoxes is necessarily the unit interval. Therefore, the paradoxes are blocked by the conditional event interpretation. This situation is investigated empirically in Section 3.3 for the first time. The data suggest, that the clear majority of people neither endorse Paradox 1 nor Paradox 2. Psychologically, this speaks for the conditional event interpretation of indicative conditionals. Future empirical research is needed to clarify the case of the paradoxes in the context of additional probabilistic or logical knowledge.

The conditional introduction inference from the affirmed consequent

[2]Cf., e. g., the *Bill Gates example*.

(Paradox 1) deserves special attention. Bonnefon and Politzer (2010, p. 154) note that

> "[…] a case of interest is that in which the premise y is certain. In that case, 'If x, y' must also be certain, and the inference is valid. This specific case remained to be accounted for by probabilistic approaches to conditionals; they can now rely on our pragmatic account."

This remark is true for the standard approach to probability, where $P(B) = 1$ implies $P(B|A) = 1$ (provided that $P(A) > 0$, otherwise $P(B|A)$ is *undefined*).[3] However, this is not true in the coherence approach to probability. I include a proof in Appendix B (see p. 134) which demonstrates that $P(B|A) \in [0,1]$ is coherent even in the very special case when $P(B) = 1$. Thus, coherence based probability logic provides a purely semantical rational reconstruction of the paradoxes without the need of superimposing pragmatic considerations.

2.4 Nonmonotonic inference

Section 3.4 (p. 76) investigates nonmonotonic reasoning from a formal and an empirical point of view. Nonmonotonic reasoning investigates formal structures governing the rational retraction of conclusions in the light of new evidence. The paradigm example is the *Tweety case*. From the two premises *Tweety is a bird* and *birds can fly* it seems valid to conclude that Tweety can fly. If the premise set is augmented by the further premise *Tweety is a penguin*, common sense tells us to retract the conclusion that Tweety can fly, as penguins usually cannot fly. In the framework of classical logic, however, augmenting the premise set of a logically valid argument

[3] In the standard approach to probability $P(B|A) =_{def.} \frac{P(A \wedge B)}{P(A)}$, if $P(A) > 0$. If $P(B) = 1$, then $P(B|A) = 1$ (if $P(A) \neq 0$, otherwise $P(B|A)$ is undefined).

cannot lead to a not logically valid argument, i. e., classical logic is *monotonic*. Nonmonotonic formalisms were developed to account for situations like the *Tweety case*.

Much of the formal work in nonmonotonic reasoning is motivated by appealing to the intuition that people reason nonmonotonically, as the following selected quotes illustrate:

> "In everyday life, however, it seems clear that we, human beings, draw sensible conclusions from what we know and that, on the face of new information, we often have to take back previous conclusions, even when the new information we gathered in no way made us want to take back our previous assumptions" (Kraus, Lehmann, & Magidor, 1990, p. 167).

> "We will argue that humans often use circumscription, and robots must too. [...] We think *circumscription* accounts for some of the successes and some of the errors of human reasoning" (McCarthy, 1977, p. 1040).

> "To formalize human commonsense reasoning something different [than classical logic] is needed. Commonsense reasoning is frequently not monotonic" (Brewka, 1991, p. 2). Moreover, "[n]onmonotonicity seems to be a fundamental aspect of human commonsense reasoning in all kinds of areas" (Brewka, 1991, p. 13).

Based on similar quotes, Pelletier and Elio (1997) argue that nonmonotonic reasoning is a genuinely psychologistic endeavor:

> "...considering how people actually do default reasoning is an important and necessary grounding for the entire enterprise of formalizing default reasoning" (Pelletier & Elio, 1997, p. 165).

> "We have claimed in this paper that, unlike classical logic, default reasoning is basically a psychologistic enterprise" (Pelletier & Elio, 1997, p. 177).

I disagree that nonmonotonic reasoning is a genuinely psychologistic endeavor. Nonmonotonic formalisms provide many psychologically fruitful intuitions and numerous empirically testable hypotheses are derivable, but this does not mean that nonmonotonic formalisms need to be justified solely on empirical grounds. There are *a priori* rationality norms for nonmonotonic reasoning. System P, for example, is a set of basic rationality postulates any system of nonmonotonic reasoning should satisfy (Kraus et al., 1990). Moreover, Gilio (2002) proposed a coherence based probability semantics for just this system. I argued that the relation between formal and empirical work is a genuinely interactive one (Pfeifer, 2011b). On the one hand, how people reason nonmonotonically can stimulate new formalisms, like System LS (Ford, 2004). On the other hand, empirical work can provide quality criteria to evaluate formal work.

Compared to the vast literature on nonmonotonic formalisms and the vast literature on the psychology of reasoning, only a few studies investigated human nonmonotonic reasoning (Benferhat, Bonnefon, & Da Silva Neves, 2005; Elio & Pelletier, 1993; Ford & Billington, 2000; Pfeifer, 2002, 2006b; Pfeifer & Kleiter, 2002, 2003, 2005, 2006b; Schurz, 2005; Vogel, 1996). Section 3.4 further extends the empirical work. Moreover, I investigate key inference rules of System P and contrast them with their respective monotonic counterparts within the coherence based probability semantics in Section 3.4. The experimental studies suggest that people endorse the System P rules and understand that the monotonic counterparts are probabilistically non-informative.

Section 3.4 shows how the above described paradigmatic *Tweety case* can be interpreted probabilistically. The formalization has been criticized by Schurz (2011), who argues that the *Tweety case* calls for the need of two measures of probability, an objective and an evidential (subjective) one. Schurz obtains a probabilistic incoherence in the formalization by probabilizing the known facts that Tweety is a bird and a penguin: $P(\textit{Tweety is a bird} \wedge \textit{Tweety is a penguin}) = 1$ (or at least greater than some threshold). This conjunction probability is the key to his incoherence result. I defend my formalization by arguing that if it is known that Tweety is a bird and a penguin, then it should be represented as a fact and not by a conjunction probability. This strategy is used in premises 2 and 3 in the formalization presented in Section 3.3 (see Section 4 in this paper). Thus, it is free of probabilistic incoherences. Moreover, my proposed distinction between factual premises and probabilistic premises allows for a clear distinction between the ontic and the epistemic dimension, respectively, in the analysis of arguments.

2.5 Properties of uncertain argument forms

A central topic in epistemology is the justification of beliefs. An important way to understand belief justification is the formalization of justifications in terms of premises and conclusions. As soon as a formalization is established, one can work out in what sense the conclusion (i. e., the *justificandum*) is warranted by the premise set (i. e., the *justificans*). As explained above (Section 2.1, p. 14), logical validity is not applicable for the rational evaluation of everyday life arguments as it neither accounts for the uncertainty of the premises nor for the defeasibility of the conclusion. Coherence based probability logic provides the necessary tools: attaching uncertainty to the premises, expressing the imprecision of the probabilistic assessment, and dealing adequately with the defeasibility of conclusions (cf., e. g., the *Tweety case* discussed in the previous section).

Section 3.5 contributes to the formalization of arguments and investigates interrelations among general properties of probabilistic argument forms: logical validity, probabilistic informativeness, and Adams' (1975) p-validity[4]. All three properties classify argument forms (see Figure 1 in Section 3.5). The conjunction of logical validity and probabilistic informativeness is necessary and sufficient for p-validity. p-validity is a probabilistic analog of logical validity. p-validity preserves high probabilities and logical validity preserves truth.

Probability logic determines the tightest coherent lower and upper probability bounds of the conclusion (see Table 2 in Section 3.5). I argue that it is more natural and informative for the justification of beliefs to focus on the conclusion probability than just on p-validity. Since p-validity of an argument $\mathcal{A}$ presupposes logical validity of the non-probabilistic version $\mathcal{A}$, p-validity excludes an important class of arguments that are probabilistically informative but not logically valid. This is illustrated in the following example.

The LD_{50} value of Dioxin is .02 milligrams per one kilogram bodyweight (Fischhoff & Kadvany, 2011, p. 48). This means that if a rat (call it Petra) with a bodyweight of half a kilogram ingests .018 milligrams of Dioxin, then the probability that the rat will die (within a specified period of time) is .9. Assume that the probability that Petra does not intake Dioxin is low. Based on these premises, consider the following instance of the probabilistic *denying of the antecedent*:

(1) $P(\text{Petra dies} \mid \text{Petra ingests .018 milligrams of Dioxin}) = .9$

(2) $P(\text{Petra does not ingest .018 milligrams of Dioxin}) = .1$

(3) $P(\text{Petra does not die}) \in [.09, .19]$

[4] An argument is p-valid if, and only if, the uncertainty of the conclusion of a valid inference cannot exceed the sum of the uncertainties of its premises, where "uncertainty of X" means $1 - P(X)$.

Intuitively, it is plausible that Petra dies (within a specified period of time), if the probability that it did not ingest Dioxin is low. *Denying of the antecedent* is not logically valid, but the probabilistic versions can provide informative conclusion probabilities. Such cases are missed by considering only p-valid arguments.

Note that the conclusion probability of the *Denying of the antecedent* becomes non-informative if the premise probabilities approach one. This reflects that this argument form is not logically valid.

Applying the measure of argument strength presented above (Section 2.1, p. 15), the strength of the argument in the *Petra example* obtains value .045, which is rather low (recall that the strongest arguments obtain value one and the weakest arguments obtain value zero). The above *Petra example* shows that even weak and not logically valid arguments can provide rational reasons to justify beliefs.

Finally I stress that the formalization of arguments requires cautiousness. While coherence based probability logic provides tools to represent fine grained differences, it is the job of the epistemologists to find out what and how it is to be formalized. Consider the following example that may be formalized in at least two completely different ways:

(A) Consider if A then B. Will $\neg A$, if $\neg B$?

One way of formalizing (A) is in terms of a probabilistic *modus tollens*; an alternative way is in terms of a probabilistic *contraposition*. Section 3.5 points out that: While both argument forms are logically valid in classical logic, they need to be distinguished sharply in probability logic: *modus tollens* is probabilistically informative but *contraposition* is probabilistically non-informative. In some psychological investigations this difference seems to remain undetected which has disastrous consequences not only for the interpretation of the data. It is also obvious that one should be aware of this difference in epistemological discourses.

Chapter 3

Collected papers

On argument strength

Niki Pfeifer
(niki.pfeifer@lrz.uni-muenchen.de)
Munich Center for Mathematical Philosophy
Ludwig-Maximilians-Universität München
Geschwister-Scholl-Platz 1, 80539 Munich, Germany

Abstract

Everyday life reasoning and argumentation is defeasible and uncertain. I present a probability logic framework to rationally reconstruct everyday life reasoning and argumentation. Coherence in the sense of De Finetti is used as the basic rationality norm. I discuss two basic classes of approaches to construct measures of argument strength. The first class imposes a probabilistic relation between the premises and the conclusion. The second class imposes a deductive relation. I argue for the second class, as the first class is problematic if the arguments involve conditionals. I present a measure of argument strength that allows for dealing explicitly with uncertain conditionals in the premise set.

Probabilistic approaches to argumentation became popular in various fields including argumentation theory (e.g., Hahn & Oaksford, 2006), formal epistemology (e.g., Pfeifer, 2007, 2008), the psychology of reasoning (e.g., Hahn & Oaksford, 2007), and computer science (e.g., Haenni, 2009).

Probabilistic approaches allow for dealing with the uncertainty and defeasibility of everyday life arguments. This paper presents a procedure to formalize everyday life arguments in probability logical terms and to measure its strength.

"Argument" denotes an ordered triple consisting of (i) a (possibly empty) premise set, (ii) a conclusion indicator (usually denoted by "therefore" or "hence"), and (iii) a conclusion. As an example consider the following argument $\mathcal{A}$:

(1) If Tweety is a bird, then Tweety can fly.

(2) Tweety is a bird.

(3) Therefore, Tweety can fly.

In terms of the propositional calculus, $\mathcal{A}$ can be represented by $\mathcal{A}_1$:

(1) $B \supset F$

(2) B

(3) $\therefore\ F$

where "B" denotes "Tweety is a bird.", "F" denotes "Tweety can fly.", "$\therefore$" denotes the conclusion indicator, and "$\supset$" denotes the material conditional. The material conditional ($A \supset B$) is false if the antecedent (A) is true and the consequent (B) is false, and true otherwise.[1]

Argument $\mathcal{A}_1$ is an instance of the logically valid *modus ponens*. An argument is logically valid if, and only if, it is impossible that all premises

[1]Note that the propositional-logically atomic formulae B and F in argument $\mathcal{A}_1$ can be represented in predicate logic by $Bird(Tweety)$ and $Can\_Fly(Tweety)$, respectively. Moreover, F may be represented even more fine-grained in modal logical terms by $\diamond F$, where "$\diamond$" denotes a possibility operator. However, for the sake of sketching a theory of argument strength, it is sufficient to formalize atomic propositions by propositional variables.

are true and the conclusion is false. In everyday life, however, premises are often uncertain and conditionals allow for exceptions. Not all birds fly: penguins, for example, are birds that do not fly. Also the second premise may be uncertain: Tweety could be a non-flying bird or not even a bird. This uncertainty and defeasibility cannot be properly expressed in the language of the propositional calculus. Nevertheless, as long as there is no evidence that Tweety is a bird that cannot fly (e.g., that Tweety is a penguin), the conclusion of $\mathcal{A}$ is rational.

Probability logic allows for dealing with exception and uncertainty (e.g., Adams, 1975; Hailperin, 1996; Coletti & Scozzafava, 2002). It provides tools to reconstruct the rationality of reasoning and argumentation in the context of arguments like $\mathcal{A}_1$. Among the various approaches to probability logic, I advocate *coherence based probability logic* for formalizing everyday life arguments (Pfeifer & Kleiter, 2006a, 2009). Coherence based probability logic combines coherence based probability theory with propositional logic. It received strong empirical support in a series of experiments on the basic nonmonotonic reasoning System P (Pfeifer & Kleiter, 2003, 2005, 2006b), the paradoxes of the material conditional (Pfeifer & Kleiter, 2011), the conditional syllogisms (Pfeifer & Kleiter, 2007), and on how people interpret (Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011) and negate conditionals (Pfeifer, 2012).

Coherence based probability theory was originated by De Finetti (1980, 1974). Among others, it has been further developed by Walley (1991), Lad (1996), Biazzo and Gilio (2000), and Coletti and Scozzafava (2002). In the framework of coherence, probabilities are (subjective) *degrees of belief* and not objective quantities. It seems natural that different people may assign different degrees of belief to the premises of one and the same argument.

This does not mean, however, that everything is subjective and therefore no general rationality norms are available. *Coherence* requires to avoid bets that lead to sure loss, which in turn guarantees that the axioms of probability theory are satisfied.[2] Another characteristic feature of coherence is that conditional probability, $P(B \mid A)$, is a *primitive* notion. Consequently, the probability value is assigned *directly* to the conditional event, $B \mid A$, as a whole. This contrasts with the standard approaches to probability, where conditional probability ($P(B \mid A)$) is defined by the fraction of the joint and the marginal probability ($P(A \wedge B) / \Pr(A)$). The probability axioms are formulated for conditional probabilities and not for absolute probabilities (the latter is done in the standard approach to probability and is problematic if $P(A) = 0$). Coherence based probability logic tells us how to propagate the uncertainty of the premises to the conclusion. As an example consider a probability logical version of the above argument, $\mathcal{A}_2$:

(1) $P(F|B) = x$
(2) $P(B) = y$
(3) $\therefore\ xy \leq P(F) \leq xy + 1 - y$

where xy and $xy + 1 - y$ are the tightest coherent lower and upper probability bounds, respectively, of the conclusion. $\mathcal{A}_2$ is an instance of the probabilistic modus ponens (see, e.g., Pfeifer & Kleiter, 2006a). If premise (1) had been replaced by the probability of the material conditional, then the tightest coherent lower and upper probability bounds of the conclusion would have been different ones. However, paradoxes and experimental results suggest that uncertain conditionals should not be represented by the

[2] I argued elsewhere (Pfeifer, 2008) that violation of coherence is a necessary condition for an argument to be fallacious.

probability of the material conditional ($P(A \supset B)$), but rather by the conditional probability ($P(B|A)$; Pfeifer & Kleiter, 2010, 2011).

The consequence relation between the premises and the conclusion is deductive in the framework of coherence based probability logic. The probabilities of the premises are transmitted deductively to the conclusion. Depending on the logical and probabilistic structure of the argument, the best possible coherent probability bounds of the conclusion can be a *precise* (point) probability value or an imprecise (interval) probability. Interval probabilities are constrained by a lower and an upper probability bound (see the conclusion of $\mathcal{A}_2$). In the worst case, the unit interval is a coherent assessment of the probability of the conclusion. In this case the argument form is probabilistically non-informative: zero and one are the tightest coherent probability bounds (Pfeifer & Kleiter, 2009, 2006a).

The tightest coherent probability bounds of the conclusion provide useful building blocks for a measure of argument strength. Averages of the tightest coherent lower and upper probabilities of the conclusion given some threshold probabilities of the premises allow for measuring the strength of *argument forms* (like the modus ponens; see Pfeifer & Kleiter, 2006a). In the following I focus on measuring the strength of *concrete arguments* (like argument $\mathcal{A}$).

There are at least two alternative ways to construct measures of argument strength: one presupposes a *deductive* consequence relation, whereas the other one presupposes an *uncertain* consequence relation. As explained above, coherence based probability logic involves a deductive consequence relation. Theories of confirmation assume that there is an uncertain relation between the evidence and the hypothesis. "Theories of confirmation may be cast in the terminology of argument strength, because $P_1 \ldots P_n$ confirm C

$S_d(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{C}\vert\mathcal{P}) - P(\mathcal{C})$	(Carnap, 1962)
$S_s(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{C}\vert\mathcal{P}) - P(\mathcal{C}\vert\neg\mathcal{P})$	(Christensen, 1999)
$S_m(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{P}\vert\mathcal{C}) - P(\mathcal{P})$	(Mortimer, 1988)
$S_n(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{P}\vert\mathcal{C}) - P(\mathcal{P}\vert\neg\mathcal{C})$	(Nozick, 1981)
$S_c(\mathcal{P},\mathcal{C})$	$=$	$P(\mathcal{P} \wedge \mathcal{C}) - P(\mathcal{P}) \times P(\mathcal{C})$	(Carnap, 1962)
$S_r(\mathcal{P},\mathcal{C})$	$=$	$\frac{P(\mathcal{C}\vert\mathcal{P})}{P(\mathcal{C})} - 1$	(Finch, 1960)
$S_g(\mathcal{P},\mathcal{C})$	$=$	$1 - \frac{P(\neg\mathcal{C}\vert\mathcal{P})}{P(\neg\mathcal{C})}$	(Rips, 2001)
$S_l(\mathcal{P},\mathcal{C})$	$=$	$\frac{P(\mathcal{P}\vert\mathcal{C}) - P(\mathcal{P}\vert\neg\mathcal{C})}{P(\mathcal{P}\vert\mathcal{C}) + P(\mathcal{P}\vert\neg\mathcal{C})}$	(Kemeny & Oppenheim, 1952)

Table 1: Measures of confirmation presented in the literature (adapted from Crupi et al., 2007).

only to the extend that $P_1 \ldots P_n / C$ is a strong argument." (Osherson, Smith, Wilkie, López, & Shafir, 1990, p. 185). Table 1 casts a number of prominent measures of confirmation in terms of argument strength.

The underlying intuition of measures of confirmation is that premise set $\mathcal{P}$ *confirms* conclusion $\mathcal{C}$, if the conditional probability of the conclusion given the premises is higher than the absolute probability of the conclusion, $P(\mathcal{C}|\mathcal{P}) > P(\mathcal{C})$. $\mathcal{P}$ *disconfirms* $\mathcal{C}$, if $P(\mathcal{C}|\mathcal{P}) < P(\mathcal{C})$. If $\mathcal{C}$ is stochastically independent of $\mathcal{P}$, i.e. $P(\mathcal{C}|\mathcal{P}) = P(\mathcal{C})$, then the premises are *neutral* w.r.t. the confirmation of the conclusion. As pointed out by Fitelson (1999), these three conditions do not impose restrictions on the choice of the measures in Table 1, i.e., they are satisfied in the context of the listed measures.

Measures of confirmation may be appropriate for measuring the strength of arguments if we do not want to formalize explicitly the structure of the premise set. However, if the premise set includes conditionals (like argument $\mathcal{A}$), then these measures require a theory of how to combine conditionals and how to conditionalize on conditionals. Consider, for example argument $\mathcal{A}$ and the general requirement that a strong argument should satisfy the inequality $P(\mathcal{C}|\mathcal{P}) > P(\mathcal{C})$. It is easy to instantiate the conclusion of $\mathcal{A}$: $P(B|\mathcal{P}) > P(B)$. There are at least two options to instantiate the premise set $\mathcal{P}$. Both options depend on how the conditional in premise 1 is interpreted.

The first option consists in the interpretation of the conditional in terms of a conditional event, $B|A$. In this case at least two problems need to be solved. The first one is the combination of the conditional premise(s) with the other premise(s): "$(B|A) \,\&\, A$" is not defined.[3] The second problem concerns the conditionalization on conditionals: the meaning of "$P(B \mid (B|A) \ldots)$" needs to be explicated. This is a deep problem and an uncontroversial general theory is still missing (for a proposal of how to conditionalize on conditionals see, e.g., Douven, 2012).

The second option consists in the interpretation of the conditional in terms of the material conditional, $A \supset B$. Here, it is straightforward to combine the material conditionals and to conditionalize on the material conditional. Argument $\mathcal{A}$ is instantiated in the general requirement of strong arguments as follows: $P(B \mid A \wedge (A \supset B)) > P(B)$. However, coherence requires that $P(B \mid A \wedge (A \supset B)) = 1$. Thus, the inequality is trivially satisfied (if $P(\mathcal{C}) < 1$). It is counterintuitive that any instance—including those with low premise probabilities—of $\mathcal{A}$ are strong arguments. Therefore, measures of confirmation are not appropriate measures of argument strength if

[3]Since the conditional event is non-propositional, it cannot be combined by classical logical conjunction. Conditional events *can* be combined by so-called "quasi-conjunctions" (Adams, 1975, p. 46 f). As Adams notes, however, quasi-conjunctions lack some important logical features of conjunctions.

we want to explicitly formalize arguments that include conditionals.

I will now turn to a measure of argument strength and show how it allows for formalizing arguments that involve conditionals. The crucial idea is that (i) the precision of a strong argument is high and that (ii) the location of the coherent probability (interval) is close to 1 (Pfeifer, 2007). The imprecision is measured by the size of the tightest coherent probability bounds of the conclusion. Let z' and z'' denote the tightest coherent lower and upper bounds, respectively, of an argument $\mathcal{A}_x$. The imprecision of $\mathcal{A}_x$ is measured by $z'' - z'$. Consequently, the *precision* of $\mathcal{A}_x$ is measured by $1 - (z'' - z')$. The location of the coherent conclusion probability is measured by the arithmetic mean of the tightest coherent probability bounds, $\frac{z' + z''}{2}$. The argument strength s of $\mathcal{A}_x$ is equal to the product of the precision and the location of the tightest coherent probability bounds of the conclusion:

$$s(\mathcal{A}_x) \;=\; [1 - (z'' - z')] \times \frac{z' + z''}{2},$$

where $0 \leq s(\mathcal{A}_x) \leq 1$, since $0 \leq z' \leq z'' \leq 1$. The values 0 and 1 denote the weakest and the strongest value, respectively.

As an example of the evaluation procedure of the strength of an argument, consider the following instance of argument $\mathcal{A}_2$:

(1) $P(F|B) = .8$

(2) $P(B) = .9$

(3) $\therefore\ .72 \leq P(F) \leq .82$

The strength of this argument is .69. In the special case where the premises are certain (i.e., probabilities equal to 1) the strength of the argument obtains its maximum value 1.

Figure 1 presents the behavior of the measure in general. According to the measure, the argument strength increases if the location of the tightest

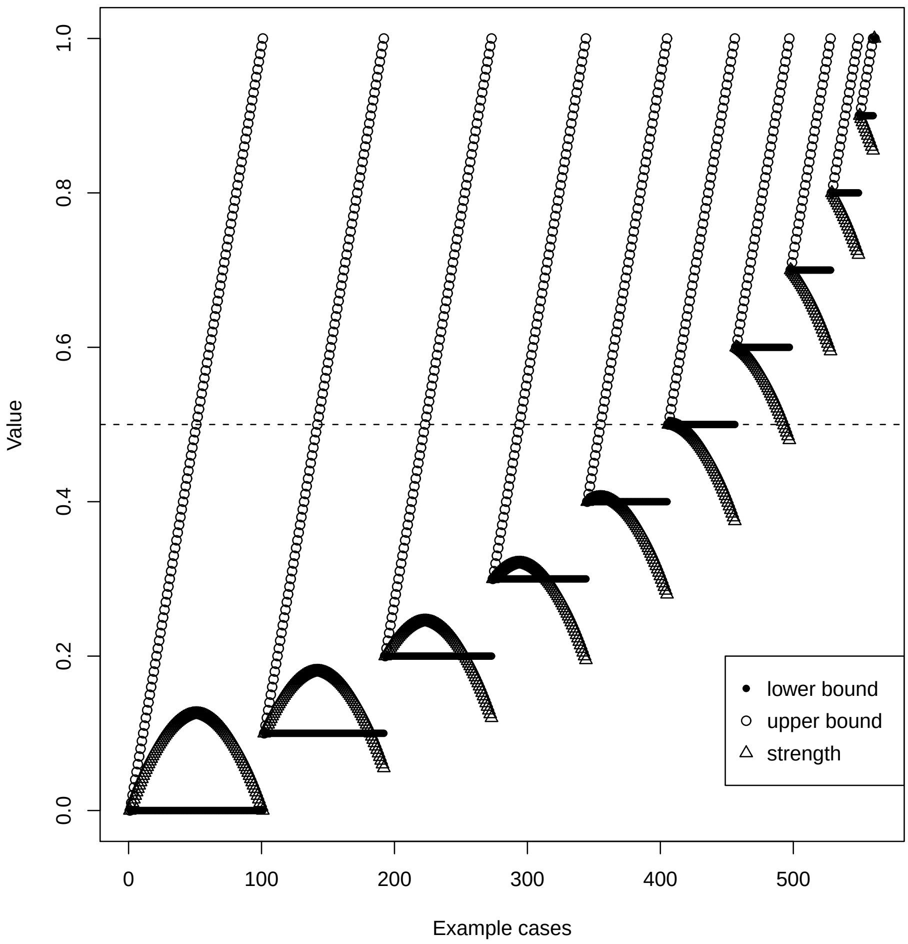

Figure 1: Let z' denote the tightest coherent lower and z'' denote the tightest coherent upper bound of an argument $\mathcal{A}$. The argument strength of $\mathcal{A}$ is equal to $[1-(z''-z')] \times \dfrac{z'+z''}{2}$. The strength of $\mathcal{A}$ increases if the precision of the conclusion is high and the location of the tightest coherent probability interval is close to 1.

coherent bounds of the conclusion approaches 1. The argument strength decreases if the imprecision increases. Moreover, an argument is weak if the conclusion probability is low. Maximum imprecision implies minimum argument strength. It follows that all probabilistically non-informative arguments are also weak arguments (with $s = 0$). Figure 2 shows the behavior of the measure for coherent lower conclusion probabilities of at least .5. If the conclusion probability is at least .5, then the argument strength varies between .375 and .500. The higher the precision the higher the strength of the argument.

The proposed measure contrasts with the traditional measures of confirmation presented in Table 1. The consequence relation remains deductive, while measures of confirmation assume an uncertain relation between the premises and the conclusion. Using probability logic to formalize arguments is advantageous as it does justice to the logical structure: premise sets that include conditionals can be represented explicitly. If a measure of argument strength requires to calculate the conditional probability of the conclusion given some combination of the premises, $P(\text{conclusion}\,|\,\text{premise set})$, then severe problems arise of how to connect premises containing conditionals with each other and how to conditionalize on conditionals. In the proposed measure this problem is avoided, as probability logic tells us how to infer the tightest coherent probability bounds of the conclusion from the premises, which are in turn exploited for calculating the argument strength.

The proposed measure s has not only attractive theoretical consequences (as explained above), it also implies at least two psychologically plausible hypothesis. People judge arguments as strong, if the premises imply high conclusion probabilities (i) and if the conclusion probability

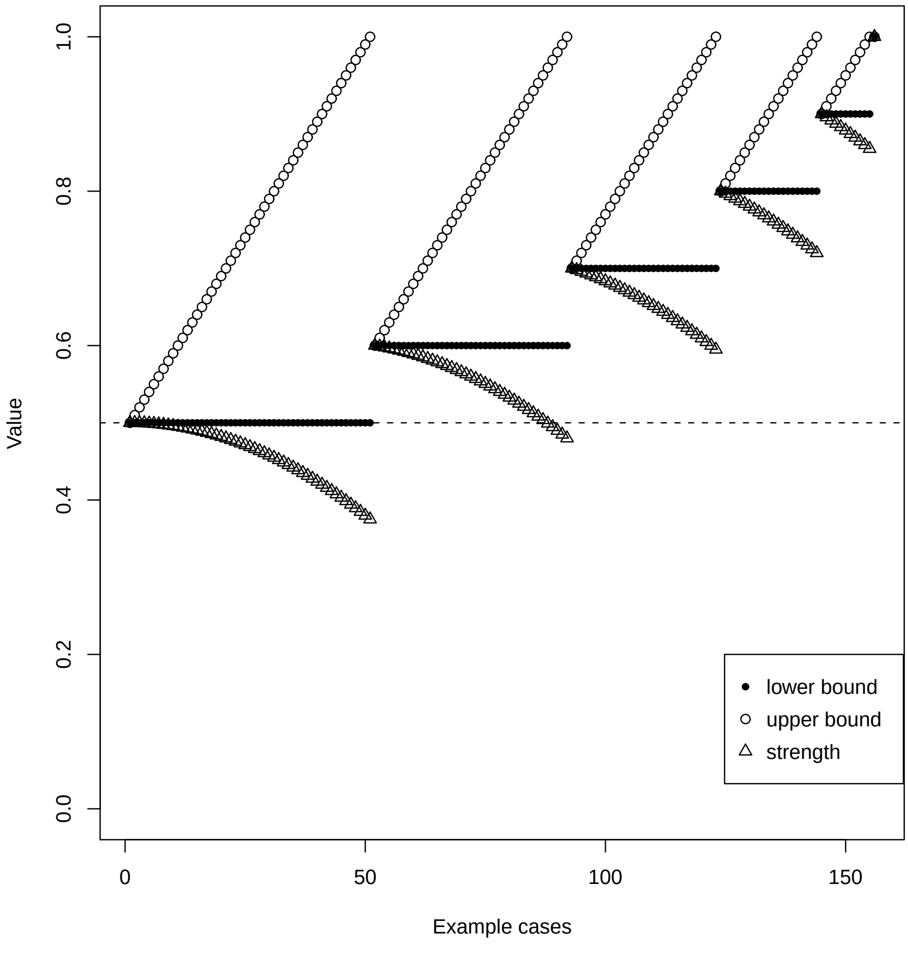

Figure 2: Detail of Figure 1, showing the behavior of measure *s* for coherent lower conclusion probabilities of at least .5.

is—at the same time—precise (ii). The empirical test of these hypothesis is a challenge for future research.

Acknowledgments

This work is financially supported by the *Alexander von Humboldt Foundation*, the *German Research Foundation* project PF 740/2-1 "Rational reasoning with conditionals and probabilities. Logical foundations and empirical evaluation" (Project leader: Niki Pfeifer; Project within the DFG Priority Program SPP 1516 "New Frameworks of Rationality"), and the *Austrian Science Fund* project P20209 "Mental probability logic" (Project leader: Niki Pfeifer).

References

Adams, E. W. (1975). *The logic of conditionals*. Dordrecht: Reidel.

Biazzo, V., & Gilio, A. (2000). A generalization of the fundamental theorem of De Finetti for imprecise conditional probability assessments. *International Journal of Approximate Reasoning*, *24*(2-3), 251-272.

Carnap, R. (1962). *Logical foundations of probability* (2nd ed.). Chicago: University of Chicago Press.

Christensen, D. (1999). Measuring confirmation. *Journal of Philosophy*, *96*, 437–461.

Coletti, G., & Scozzafava, R. (2002). *Probabilistic logic in a coherent setting*. Dordrecht: Kluwer.

Crupi, V., Tentori, K., & Gonzales, M. (2007). On Bayesian measures of confirmation. *Philosophy of Science*, *74*, 229-252.

De Finetti, B. (1974). *Theory of probability* (Vols. 1, 2). Chichester: John Wiley & Sons. (Original work published 1970)

De Finetti, B. (1980). Foresight: Its logical laws, its subjective sources (1937). In H. J. Kyburg & H. E. Smokler (Eds.), *Studies in subjective probability* (p. 55-118). Huntington, New York: Robert E. Krieger Publishing Company.

Douven, I. (2012). Learning conditional information. *Mind & Language*, *27*(3), 239–263.

Finch, H. A. (1960). Confirming power of observations metricized for decisions among hypotheses. *Philosophy of Science*, *27*, 293–207 (part I), 391–404 (part II).

Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science*, *66*, 362–378.

Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts towards the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 635–648.

Haenni, R. (2009). Probabilistic argumentation. *Journal of Applied Logic*, 155–176.

Hahn, U., & Oaksford, M. (2006). A normative theory of argument strength. *Informal Logic*, *26*, 1-22.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, *114*(3), 704-732.

Hailperin, T. (1996). *Sentential probability logic. Origins, development, current status, and technical applications*. Bethlehem: Lehigh University Press.

Kemeny, J., & Oppenheim, P. (1952). Degrees of factual support. *Philosophy* *of Science*, *19*, 307–324.

Lad, F. (1996). *Operational subjective statistical methods: A mathematical, philosophical, and historical introduction*. New York: Wiley.

Mortimer, H. (1988). *The logic of induction*. Paramus, NJ: Prentice Hall.

Nozick, R. (1981). *Philosophical explanations*. Oxford: Clarendon.

Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*(2), 185–200.

Pfeifer, N. (2007). Rational argumentation under uncertainty. In G. Kreuzbauer, N. Gratzl, & E. Hiebl (Eds.), *Persuasion und Wissenschaft: Aktuelle Fragestellungen von Rhetorik und Argumentationstheorie* (p. 181-191). Wien: LIT.

Pfeifer, N. (2008). A probability logical interpretation of fallacies. In G. Kreuzbauer, N. Gratzl, & E. Hiebl (Eds.), *Rhetorische Wissenschaft: Rede und Argumentation in Theorie und Praxis* (pp. 225–244). Wien: LIT.

Pfeifer, N. (2012). Experiments on Aristotle's Thesis: Towards an experimental philosophy of conditionals. *The Monist*, *95*(2), 223–240.

Pfeifer, N., & Kleiter, G. D. (2003). Nonmonotonicity and human probabilistic reasoning. In *Proceedings of the 6th workshop on uncertainty processing* (p. 221-234). Hejnice: September 24–27$^{\text{th}}$, 2003.

Pfeifer, N., & Kleiter, G. D. (2005). Coherence and nonmonotonicity in human reasoning. *Synthese*, *146*(1-2), 93-109.

Pfeifer, N., & Kleiter, G. D. (2006a). Inference in conditional probability logic. *Kybernetika*, *42*, 391-404.

Pfeifer, N., & Kleiter, G. D. (2006b). Is human reasoning about nonmonotonic conditionals probabilistically coherent? In *Proceedings of the 7th workshop on uncertainty processing* (p. 138-150). Mikulov: September 16–20$^{\text{th}}$, 2006.

Pfeifer, N., & Kleiter, G. D. (2007). Human reasoning with imprecise probabilities: Modus ponens and Denying the antecedent. In G. De Cooman, J. Vejnarová, & M. Zaffalon (Eds.), *5th International Symposium on Imprecise Probability: Theories and Applications* (p. 347-356). Prague, Czech Republic: SIPTA.

Pfeifer, N., & Kleiter, G. D. (2009). Framing human inference by coherence based probability logic. *Journal of Applied Logic*, *7*(2), 206–217.

Pfeifer, N., & Kleiter, G. D. (2010). The conditional in mental probability logic. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thought* (p. 153-173). Oxford: Oxford University Press.

Pfeifer, N., & Kleiter, G. D. (2011). Uncertain deductive reasoning. In K. Manktelow, D. E. Over, & S. Elqayam (Eds.), *The science of reason: A Festschrift for Jonathan St B.T. Evans* (p. 145-166). Hove: Psychology Press.

Rips, L. J. (2001). Two kinds of reasoning. *Psychological Science*, *12*(2), 129-134.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. London: Chapman and Hall.

Experiments on Aristotle's Thesis: Towards an Experimental Philosophy of Conditionals

Abstract

Two experiments (N_1 = 141, N_2 = 40) investigate two versions of Aristotle's Thesis for the first time. Aristotle's Thesis is a negated conditional, which consists of one propositional variable with a negation either in the antecedent (version 1) or in the consequent (version 2). This task allows us to infer if people interpret indicative conditionals as material conditionals or as conditional events. In the first experiment I investigate between-participants the two versions of Aristotle's Thesis crossed with abstract versus concrete task material. The modal response for all four groups is consistent with the conditional event and inconsistent with the material conditional interpretation. This observation is replicated in the second experiment. Moreover, the second experiment rules out scope ambiguities of the negation of conditionals. Both experiments provide new evidence against the material conditional interpretation of conditionals and support the conditional event interpretation. Finally, I discuss implications for modeling indicative conditionals and the relevance of this work for experimental philosophy.

1. Introduction

Experimental philosophers have investigated people's intuitions on a wide variety of philosophical topics, including causation, consciousness, cross-cultural intuitions, epistemology, morality, free will, and intentional action (see, e.g., Phillips 2011, Knobe and Nichols 2008, Feltz 2009). Conditionals, however, have not yet been discussed by experimental philosophers. This paper aims to extend the domain of experimental philosophy to conditionals. After a brief review of philosophical and psychological intuitions on probabilistic interpretation of indicative conditionals,

"Experiments on Aristotle's Thesis:
Towards an Experimental Philosophy of Conditionals" by Niki Pfeifer,
The Monist, vol. 95, no. 2, pp. 223–240.

I report two new experiments on reasoning about conditionals to clarify the interpretation and negation of indicative conditionals.

There is a long tradition of psychological investigations of conditionals. Standard tasks include Wason's selection task, truth table tasks, and conditional elimination tasks (like modus ponens; see, e.g., Evans, Newstead, and Byrne 1993). The propositional calculus was taken for granted as *the* rationality norm: rational inferences are consistent with the laws of logic and indicative conditionals (of the form "If A, then B") should be interpreted as material conditionals (denoted by $A \supset B$). People's inferences, however, diverged from the rationality postulates of classical logic.

Philosophers argued for probabilistic interpretations of indicative conditionals by relating conditionals to conditional probability, $P(B|A)$ (e.g., Adams 1975, Bennett 2003, Douven 2008, Ramsey 1978). The argument of the conditional probability function is the conditional event $B|A$. The conditional event cannot be captured within the framework of classical logic. Contrary to the material conditional ($A \supset B$) and the conjunction ($A \wedge B$), the conditional event cannot be expressed by any Boolean function: the conditional event is *void*, if the antecedent (A) is false (see Table 1)

A	*B*	*Material conditional* $A \supset B$	*Conjunction* $A \wedge B$	*Conditional event* $B\|A$
true	*true*	*true*	*true*	*true*
true	*false*	*false*	*false*	*false*
false	*true*	*true*	*false*	*void*
false	*false*	*true*	*false*	*void*

Table 1: The three most prominent psychological predictions of the interpretation of indicative conditionals. They differ only in the last two lines, where the antecedent (A) is false. The conditional event is *partially* truth-functional, as it is *void*, if the antecedence is false.

There are striking a priori reasons for why indicative conditionals are not material ones. As an example, imagine that Jones "is about to be dealt a five card poker hand from a shuffled deck of 52 cards" (Adams 2005, 1). Someone asserts:

If Jones's first card is an ace, then Jones's second card is an ace.

How sure can you be, that the sentence in the box holds? If you interpret the sentence as a conditional event and assign a conditional probability, you obtain a very low probability value,[1] which is intuitively plausible,

P(Jones's second card is an ace | Jones's first card is an ace) = 3/51 ≅ 0.06.

If you interpret the sentence in the box as a material conditional, you obtain in this scenario an intuitively implausible high probability value,[2]

P(Jones's first card is an ace ⊃ Jones's second card is an ace) = 205/221 ≅ 0.93.

Some of the arguments in favor of the conditional event interpretation have been confirmed empirically. Contrary to the material conditional, the conditional event interpretation avoids the paradoxes of the material conditional and people do not endorse these paradoxes (Pfeifer and Kleiter 2011). Premise strengthening and contraposition do not hold under the conditional event interpretation and people do not endorse these argument forms (Pfeifer and Kleiter 2010).

In recent years, probabilistic rationality norms emerged in the psychology of reasoning to better deal with the defeasibility and uncertainty and to match closer everyday inference (e.g., Evans and Over 2004, Oaksford and Chater 2009, Pfeifer and Kleiter 2005b; 2009). Within these probabilistic approaches, a new hypothesis concerning the interpretation of conditionals emerged: people interpret indicative conditionals (If A, then B) as conditional events ($B|A$). Recent studies (e.g., Fugard, Pfeifer, and Mayerhofer 2011, Fugard, Pfeifer, Mayerhofer, and Kleiter 2011, Oberauer 2006, Over, Hadjichristidis, Evans, Handley, and Sloman 2007) provide strong empirical evidence for this hypothesis.

The present paper further investigates the conditional event hypothesis. Specifically, it investigates two versions of Aristotle's Thesis for the first time. Aristotle's Thesis is a negated conditional, which consists of one propositional variable with a negation either in the antecedent (version 1) or in the consequent (version 2):

(AT #1) $\neg(\neg A \rightarrow A)$

(AT #2) $\neg(A \rightarrow \neg A)$

"$\neg$" denotes negation.[3] "$A \to B$" denotes the indicative conditional *If A, then B*, where the semantics of $\to$ is not specified. AT #1 and AT #2 are intuitively plausible. Consider an instance in natural language:

It is not the case that: If I do not win the lottery, then I win the lottery.

Likewise, it is plausible to assert

It is not the case that: If I win the lottery, then I do not win the lottery.

In classical logic, where "$\to$" is interpreted as a material conditional, (AT #1) and (AT #2) are not theorems. Both formulas are contingent ("$\equiv$" denotes equivalence):

$$\neg(\neg A \supset A) \equiv \neg A \wedge \neg A \equiv \neg A$$

$$\neg(A \supset \neg A) \equiv A \wedge A \equiv A$$

In this paper, I propose an interpretation that justifies the rationality of high beliefs in AT #1 and AT #2. Specifically, I formalize AT in terms of coherence based probability logic (Pfeifer and Kleiter 2006a; 2009).

In the psychology of reasoning three interpretations of indicative conditionals (If A, then B) are currently debated: the conditional event interpretation ($P(B|A)$), the material conditional interpretation ($P(A \supset B)$), and the conjunction interpretation ($P(A \wedge B)$). The conjunction interpretation is discussed in the theory of mental models (Johnson-Laird and Byrne 2002) and the suppositional theory of conditional reasoning (Evans and Over 2004). Both theories predict, roughly speaking, that if people process conditionals superficially, then they use the conjunction interpretation.

The coherence approach to probability goes back to De Finetti (1980, 1974) and more recent work includes, e.g., Walley (1991), Lad (1996), Biazzo and Gilio (2000), and Coletti and Scozzafava (2002). Coherence is in the tradition of subjective probability theory in which probabilities are conceived as *degrees of belief*. Degrees of belief are coherent descriptions of incomplete knowledge states. One key feature is that the coherence approach defines the probability function on an *arbitrary* family of conditional events. Therefore, it does not require a complete algebra as in the standard approach to probability. Conditional probability, $P(B|A)$, is a *primitive* notion. The probability value is assigned *directly* to the conditional event, $B|A$, as a whole (and not by definition via the fraction of the joint and the mar- ginal probability, $P(A \wedge B)/P(A)$). Therefore, the probability axioms are formulated for conditional probabilities in the framework of coherence and not for absolute probabilities (as it is done in the standard approach to probability).

Coherence based probability logic defines the consequence relation as a *deductive* one. The probabilistic inference problem consists of how to transmit the probabilities of the premises to the probability of the conclusion. Usually, the coherent probability of the conclusion is constrained by a lower and an upper probability. The coherent conclusion of the modus ponens, for example, is in the interval $xy \leq P(B) \leq xy + 1 - x$, where the two premises are $P(A) = x$ and $P(B|A) = y$, respectively.[4] As only two probabilities are given, the coherent probability of the conclusion is *imprecise*. If $P(B|\neg A) = z$ is added to the premise set of the probabilistic modus ponens, then the resulting argument form allows for inferring a *precise* probability value of the conclusion: $P(B) = xy + (1 - x)z$.

Coherence based probability logic (Pfeifer and Kleiter 2006a; 2009) has received strong empirical support in a series of experiments on the rules of the nonmonotonic System P (Pfeifer and Kleiter 2003; 2005a; 2006b), the paradoxes of the material conditional (Pfeifer and Kleiter 2011), the conditional syllogisms (Pfeifer and Kleiter, 2007), and on how people interpret conditionals (Fugard, Pfeifer, and Mayerhofer 2011; Fugard, Pfeifer, Mayerhofer, and Kleitner 2011).

Table 2 lists the probability logical predictions according to the different interpretations of indicative conditionals. The conditional event and the conjunction interpretation predict that people should hold a strong belief in both versions of AT: the probability value 1 is the only coherent assessment. The material conditional interpretation predicts that people cannot tell whether AT holds: any (point or interval) value from zero to one is coherent. Experiment 1 investigates these predictions empirically.

AT	$P(\cdot\supset\cdot)$	$P(\cdot\wedge\cdot)$	$P(\cdot|\cdot)$
$\neg(\neg A\rightarrow A)$	$0\leq P(\neg A)\leq 1$	$P(\neg(\neg A\wedge A))=1$	$P(\neg A|\neg A)=1$
$\neg(A\rightarrow\neg A)$	$0\leq P(A)\leq 1$	$P(\neg(A\supset\neg A))=1$	$P(\neg\neg A|A)=1$

Table 2: Probability logical interpretations of the two versions of Aristotle's Thesis (AT) in terms of the material conditional ($P(\cdot\supset\cdot)$), the conjunction ($P(\cdot\wedge\cdot)$), and the conditional event ($P(\cdot|\cdot)$).

2. Experiment 1

2.1 Method and design

The sample consists of 141 psychology students (110 females and 31 males). The median age of the sample is 21 (1st Qu. = 20, 3rd Qu. = 23). 91% of the participants were in their third semester.

The data were collected in a lecture hall during an introductory course on cognitive neuroscience. One week before the experiment, the students learned classical truth tables, including the material conditional and the concepts of logical truth, logical falsehood and logical contingency. At the very beginning of the unit, in which the experiment took place, the truth tables as well as related concepts were repeated. Then the lecture continued with an unrelated topic. In the middle of the lecture unit, the experiment started.

Four versions of the task material were distributed in such a way that the seating-distance between the participants of each condition was maximized. The four conditions consisted in an abstract and in a concrete version of AT #1 and AT #2.

The abstract version of AT #1 was formulated as follows:

> The letter "*A*" denotes a sentence, like "It is raining."
>
> There are sentences, where you can infer only on the basis of their logical form, whether they are guaranteed to be false or guaranteed to be true. For example:
>
> "*A* and not-*A*" is guaranteed to be false.
>
> "*A* or not-*A*" is guaranteed to be true.
>
> There are sentences, where you cannot infer only on the basis of their logical form, whether they are true or false. The sentence "*A*" ("It is raining."), for example, can be true but it can just as well be false: this depends upon whether it is actually raining.

In this part of the instruction the concepts "logical truth," "logical falsehood," and "logical contingency" are explained once again to the participants, without mentioning these technical terms explicitly. To make it clear that the task concerns the natural language version of AT as a whole and to avoid ambiguities of the scope of the question, AT was put into a box:

Evaluate the following sentence (Please tick exactly one alternative):

It is not the case, that: If not-*A*, then *A*.

The sentence in the box is guaranteed to be false	☐
The sentence in the box is guaranteed to be true	☐
One cannot infer whether the sentence is true or false	☐

The abstract version of AT #2 was identical to the abstract version of AT #1 except for the sentence in the box: the conditional "If not-*A*, then *A*" was replaced by "If *A*, then not-*A*."

The concrete version of AT was formulated by replacing the abstract letters by concrete objects. A further difference to the abstract version was the use of an implicit negation in the conditional. Negations are hard to process in general. Moreover, concrete task material is easier to process than abstract material. Thus, the concrete versions were hypothesized to be easier to process for the participants than the abstract versions.

In the concrete versions of AT, the participants were asked to imagine that there is either a dog or a cat behind a door, but not both. As in the abstract version of the task, the concepts "logical truth," "logical falsehood," and "logical contingency" were introduced informally. Then, the participants were asked to

Evaluate the following sentence (Please tick exactly one alternative):

It is not the case, that: If there is a cat behind the door, then there is a dog behind the door.

The response format was the same as in the abstract versions. The other concrete version of AT differed only in one respect: the words "cat" and "dog" changed their positions in the conditional in the box.

Strictly speaking, there is no difference between AT #1 and AT #2 in the concrete version of the task. The vignette makes clear that "cat" means "*not*-dog" and "dog" means "*not*-cat."

In all versions of the task, the task was presented together with the instructions on one page. This helps to minimize working memory demands.

2.2 Results and discussion

In all four conditions of the experiment the participants were asked to rank on a scale how clear and comprehensible the task was to them. Furthermore, the participants evaluated the confidence in the correctness of their solution and the task difficulty. Figure 1 presents the results of the participants' task evaluations. The mean rating of the task comprehensibility is close to "very clear," which indicates that the participants were not swamped by processing the conditional and the two negations. Moreover, the mean subjective confidence in the correctness of the participants' responses and the task difficulty were in the middle regions of the respective scales. This suggests that the participants did not opt out of the task. If a task is obscure or if it is perceived to be too easy or too hard, then there is a danger that the participants do not engage themselves properly in the task, and—in the worst case—opt out of doing the task. In sum, the results suggest that the task and the vignette stories are not obscure, and are perceived as being neither too easy nor too difficult.

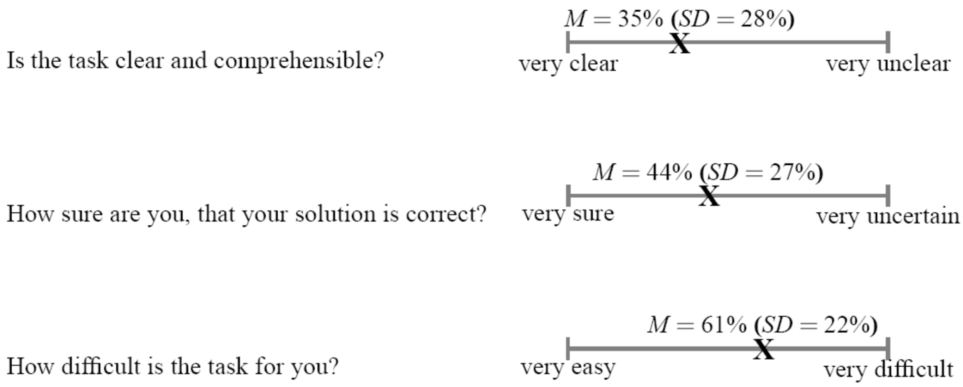

Figure 1: Mean ratings of the task comprehensibility, confidence in correctness, and difficulty (Experiment 1, N_1=141).

There are also no statistically significant differences between the responses in AT #1 and AT #2. Therefore, the data of AT #1 and AT #2 are pooled. Figure 2 summarizes the main results of Experiment 1. The modal response is consistent with the conditional event interpretation and inconsistent with the material conditional interpretation of conditionals. There is a slightly higher proportion of participants in the concrete than in the abstract condition who hold a strong belief in AT. However, this difference is statistically not significant.

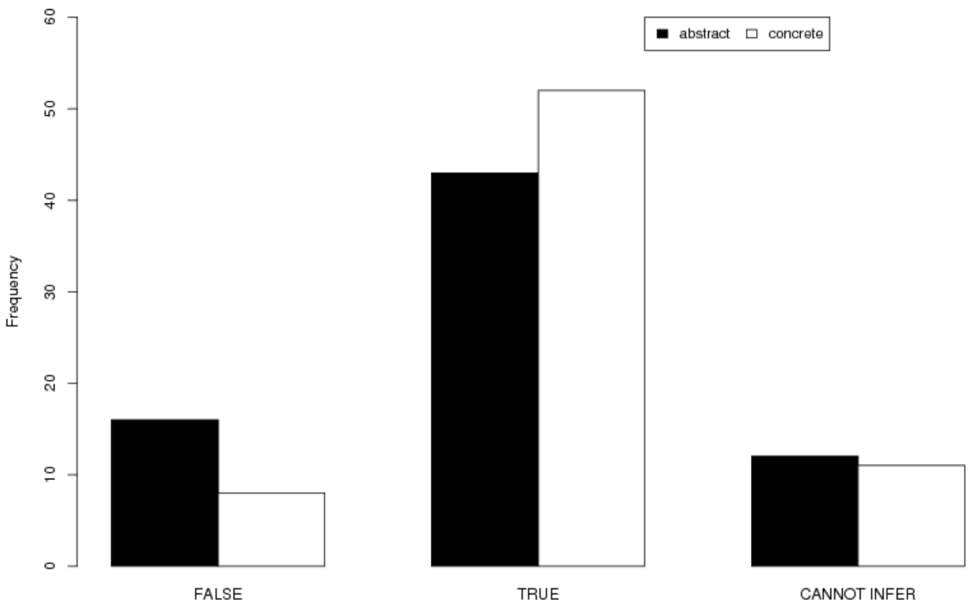

Figure 2: Response frequencies in Experiment 1 (pooled data of AT #1 and #2; N_1=141).

A minority of the participants expressed a strong belief that the sentence in the box is false. Exploration questions[5] after the experiment revealed that some participants reasoned about the conditional only. These participants overlooked that the conditional is negated. If this negation is ignored, these ratings are perfectly rational under the conditional event interpretation: coherence requires $P(A|\neg A) = 0$ and $P(\neg A|A) = 0$.

One key result of Experiment 1 is that the conditional event interpretation predicts the modal response of the participants. Another key result indicates that the literal formalization of AT by the material conditional interpretation does not predict the modal response of the participants. The data suggest that people do not interpret AT #1 as $\neg(\neg A \supset A)$ and that they do not interpret AT #2 as $\neg(A \supset \neg A)$. Thus, AT provides a watershed to experimentally differentiate between the material conditional and the conditional event interpretations of indicative conditionals. However, as noted above, AT alone does not distinguish between the conditional event and the conjunction interpretation: both predict that AT is guaranteed to be true. Experiment 2 will address this issue.

Proponents of the material conditional interpretation may argue that the results of Experiment 1 provide evidence against the wide scope reading of the negation of conditionals,

$$\underbrace{\neg(A \rightarrow \neg A)}_{\text{wide scope}}.$$

They may further argue that people interpret indicative conditionals as material conditionals and material conditionals are negated by negating the consequent,

$$(A \rightarrow \underbrace{\neg\neg}_{\text{narrow scope}} A).$$

Consequently, AT #1 reduces to $\neg A \supset \neg A$ and AT #2 reduces to $A \supset A$, which leads to the same predictions as given by the conditional event interpretation.[6]

Experiment 2 is designed to (i) clarify this scope ambiguity, (ii) differentiate between the conjunction and the conditional event interpretation of indicative conditionals, and to (iii) improve the experimental conditions.

3. Experiment 2

3.1 Method and design

Forty students (20 females and 20 males) of the University of Salzburg were tested individually in experimental rooms of the Psychology Department. Psychology students and students with a formal background were not included in the sample. The participants received 5 € for participation. Between participant explicit (n_1 = 20) versus implicit negation (n_1 = 20) conditions were varied and each participant had to solve 12 tasks (see Table 3). All tasks were concrete.

As in Experiment 1, the first part of the instruction explains the concepts "logical truth," "logical falsehood," and "logical contingency" without mentioning technical terms. In the explicit negation condition, the participants were asked to imagine the following situation:

> Hans expects to be visited by Thea and Ida. He is sitting in his room. Suddenly someone knocks at the door. Hans is absolutely certain, that either Thea or Ida is knocking.

Evaluate the following sentence (please tick exactly one alternative):

It is not the case, that: **If** Ida knocks, **then** Ida **does not** knock.

The response format was identical to the one used in Experiment 1. The logical structure of the sentence in the box was formatted in boldface to reduce the probability that participants overlook the negation in front of the conditional. Since the aim of the experiment is to investigate intuitions about the *degrees of belief* in (negated) conditionals, an epistemic component ("Hans is absolutely certain…") is added in Experiment 2. In the implicit versions of AT, "Ida *does not* knock" was replaced by "Thea knocks."

The vignette stories were adapted to other argument forms to differentiate between the conditional event and the conjunction interpretation. Table 3 lists the order of the task items in both conditions. Items #1 and #3 are designed to replicate the findings of AT #2 and AT #1, respectively, of Experiment 1. Item #2 may be called "negated reflexivity." It differentiates between the narrow and the wide scope reading of the negation of the material conditional. Item #4 ("reflexivity") differentiates between the conditional event interpretation and the conjunction interpretation of indicative conditionals. Items #5 and #6 are control items.

Item	Argument form	Prediction				Responses in percent		
			WS	NS				
		$\cdot\vert\cdot$	$\cdot\supset\cdot$	$\cdot\supset\cdot$	$\cdot\wedge\cdot$	T	F	CT/U
#1	$\neg(A\rightarrow\neg A)$	T	CT	T	T	**78**	18	5
#2	$\neg(A\rightarrow A)$	F	F	CT	CT	10	**88**	2
#3	$\neg(\neg A\rightarrow A)$	T	CT	T	T	**80**	13	8
#4	$A\rightarrow A$	T	T	T	CT	**93**	3	5
#5	$A\rightarrow B$	CT	CT	CT	CT	0	13	**88**
#6	$\neg(A\rightarrow B)$	CT	CT	CT	CT	20	3	**78**
#11	from B infer $A\rightarrow B$	U	H		U	40	0	**60**
#12	from B infer $A\rightarrow\neg B$	U	H		L	5	30	**65**

Table 3: Results ($N_2 = 40$) of Experiment 2. WS = wide and NS = narrow scope reading of negated material conditionals, CT = can't tell, T = true, F = false, U = uninformative conclusion probability, H = high conclusion probability, L = low conclusion probability. Items #7–#10 are part of the probabilistic truth table task (see text and Figure 3). Conditional event is the best predictor (**bold**).

The items #7–#10 are adapted from Pfeifer and Kleiter (2011). They serve (i) to replicate findings and (ii) to provide further possibilities to differentiate among the material conditional, conditional event, and conjunction interpretation. Items #7–#10 correspond to a version of the probabilistic truth table task, where the participants are instructed to imagine a pack of 120 cards. On each card, there is either a circle or a square, either in red or in blue. The pack consists of 40 red circle cards, 40 red square cards, 20 blue circle cards and 20 blue square cards. The pack is shuffled and then one card is randomly chosen. One cannot see what is printed on this card. The task consists in evaluating four conditionals on a scale with the labels "does not hold for sure" and "holds for sure." The four conditionals and the probability logical predictions are contained in Figure 3. The participants' interpretation can be inferred from their degree of belief in the conditional. Item #11 and #12 correspond to two paradoxes of the material conditional.

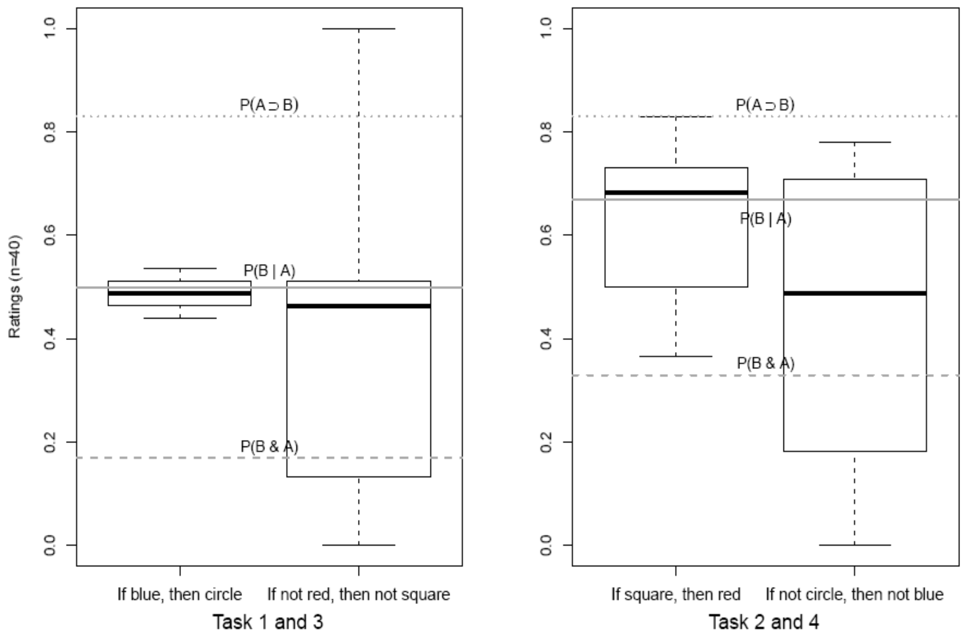

Figure 3: Results of the probabilistic truth table task (Experiment 2, $N_2 = 40$). The boxes contain 50% of the responses, and the thick line indicates the median. The whiskers indicate 1.5 × the interquartile range. Normative predictions are printed in gray (dotted, solid, and dashed lines). Conditional probability is the best predictor.

3.2 Results and discussion

There is no statistically significant difference between the two between-participant conditions. Therefore, the subsequent analysis is conducted on the pooled data. Experiment 2 replicates the results of Experiment 1: the data are consistent with the conditional event interpretation and inconsistent with the material conditional interpretation of indicative conditionals (see Table 3). Moreover, negated reflexivity (item #2) rules out the narrow scope reading of the material conditional. The conjunction interpretation is ruled out by reflexivity (item #4), the paradox of the material conditional (item #12), and the results of the probabilistic truth table task (items #7–#10).

Interpretation	Scope	Mean	SD
$\cdot\vert\cdot$		4.6	1.2
$\cdot\supset\cdot$	narrow	3.0	0.9
$\cdot\wedge\cdot$		2.6	0.9
$\cdot\supset\cdot$	wide	2.4	0.8

Table 4: Scoring of the items #1–#4 and items #11–#12 ($N_2 = 40$, 6 tasks, min = 0, max = 6). The conditional event obtains the highest score.

Scoring the data in the six tasks reveals that the mean consistency of the responses with the conditional event was the highest one out of the four interpretations (see Table 4). This reflects again that the conditional event is the best predictor for the conditional event responses.

4. Concluding remarks

Mental probability logic (Pfeifer and Kleiter 2005b; 2009), the suppositional theory of conditional reasoning (Evans and Over 2004), and the probabilistic approach by Oaksford and Chater (2009) are examples of recent psychological theories that argue for the conditional event interpretation of indicative conditionals. This study is in line with this research and provides new evidence for the conditional event interpretation.

This paper investigates Aristotle's Thesis, reflexivity and negated reflexivity, for the first time empirically. Moreover, the second experiment resolves scope ambiguities of the negation of conditionals. Neither people without training in logic (Experiment 2) nor people who just learned the truth tables and the material conditional (Experiment 1) interpret conditionals as material conditionals. The modal response pattern in all tasks corresponds to the conditional event interpretation of conditionals.

The truth functions of the material conditional and the conjunction correspond to the truth conditions of the explicit and the implicit mental models, respectively, of basic conditionals. According to the theory of mental models the core meaning of indicative conditionals is the material conditional (Johnson-Laird and Byrne 2002). The present data do not support this approach.

The tasks on Aristotle's theses differ in several respects to previous studies on conditional reasoning. First, the argument form is an inference from the empty premise set. In the abstract version, the belief in the negated conditional may be established by reasoning about a sentence in the box (i.e., the conclusion) only. In the concrete versions, the information communicated in the instructions before the box does not belong to the premise set. This information explains the relationship between the cat and the dog in this scenario. Thus, the logical form corresponds to an inference from the empty premise set as well. Second, Aristotle's thesis consists of only one propositional variable (A). However, the logical form is complex, since it is composed of two negations and one conditional. Wason's selection task, the tasks related to conditional elimination inferences (like the suppression tasks, modus ponens, modus tollens, etc.), and the variants of truth table tasks are not inferences from the empty premise set and usually involve two propositional variables.

The narrow and the wide scope readings of the negation of a conditional *If A, then B* are well defined for material conditionals ($A \supset \neg B$ and $\neg(A \supset B)$, respectively). The negation of a conditional event $B|A$, however, is well defined for the narrow scope reading only ($\neg B|A$). One might propose that negating a conditional event means that one is completely uncertain about $B|A$. Using imprecise probabilities, this could mean that the $P(B|A)$ is probabilistically uninformative, i.e., $0 \leq P(B|A) \leq 1$ is coherent. Assigning the unit interval expresses a situation of complete ignorance about $B|A$. However, the present data do not support this hypothesis: for almost all participants Aristotle's thesis is probabilistically informative.

The data of both experiments show that the acceptability/assertability conditions of $A \rightarrow B$ are consistent with $P(B|A)$ but inconsistent with

$P(A \supset B)$. Some philosophers (e.g., Lewis 1976, Grice 1975) claim that conditionals are truth functional and that $A \rightarrow B$ is acceptable/assertable *iff* (i) $P(A \supset B)$ is high, and (ii) $P(A \supset B|A)$ is high (and close to $P(A \supset B)$). On the first sight, this could be a way to save the material conditional interpretation of indicative conditionals. This is not the case: Jackson (1987, 31) notes that condition (i) and (ii) imply that $P(B|A)$ is high.

Connexive logicians investigate a branch of nonclassical logic where a standard logical vocabulary is used but certain non-theorems of classical logic like AT #1 and AT #2 are theorems (McCall 1966, Angell 2002). An implication that satisfies Aristotle's theses and the Boethius' theses $((A \rightarrow B) \rightarrow \neg(A \rightarrow \neg B)$ and $(A \rightarrow \neg B) \rightarrow \neg\,(A \rightarrow B))$ and where $\rightarrow$ cannot be understood as a biconditional (i.e., $(A \rightarrow B) \rightarrow (B \rightarrow A))$ is not a theorem) is called a *connexive implication* (Wansing 2010).

Aristotle's thesis provides an important empirical watershed between the conditional event and the material conditional interpretation of indica-tive conditionals. Other empirically interesting argument forms that allow for investigating different interpretations of conditionals include the paradoxes of the material conditional (Pfeifer and Kleiter 2011) and (non)monotonic argument forms (Pfeifer and Kleiter 2005a; 2009; 2010). The results point from different angles in the same direction: people's intuitions on the meaning of conditionals converge on the conditional event interpretation.

The present paper provides a new formalization of Aristotle's thesis in probability logical terms. Its main empirical result is that coherent conditional probabilities are natural building blocks for modeling indicative conditionals. I am convinced that "armchair philosophy" and careful experimental work can fruitfully interact. On the one hand, formal philosophy provides tools to make psychological hypotheses precise: without a proper formalism many fruitful hypotheses cannot even be formulated. On the other hand, experimental studies can empirically validate philosophical theories. Empirical investigations provide important external quality criteria for logical theories which are beyond the purely formal ones (like soundness or completeness). The present study illustrates how the domain of experimental philosophy is extended to conditionals.[7]

Niki Pfeifer

Munich Center for Mathematical Philosophy,
Ludwig-Maximilians-Universität München

Notes

1. Of all 52 cards only four ones are aces. If the first card is an ace, three aces are left for drawing a second ace out of the 51 remaining cards. Thus, P(Jones's second card is an ace | Jones's first card is an ace)=3/51 ≅ 0.06.

2. Let "A" denote the antecedent and "B" denote the consequent of the conditional. $P(A \supset B) = 1 - P(A \wedge \neg B) = 1 - (4/52 \times 48/51) = 205/221 \cong 0.93$.

3. The name "Aristotle's Thesis" was coined by McCall (1966). Aristotle wrote in his *Prior Analytics* "…if B is not great, B itself is great. But this is impossible" (quoted after Lukasiewicz, 1957, 50).

4. The law of total probability states that $P(B) = P(B|A) \times P(A) + P(B|\neg A) \times P(\neg A)$. As $P(\neg A) = 1 - P(A)$, the only unknown value is $P(B|\neg A)$. If $P(B|\neg A) = 0$, then $P(B) = P(B|A) \times P(A)$ (which is the tightest coherent lower probability bound of the conclusion). If $P(B|\neg A) = 1$, then $P(B) = P(B|A) \times P(A) + 1 - P(A)$ (which is the tightest coherent upper probability bound of the conclusion).

5. The participants were asked informally in the lecture hall how they understood the task material and how they solved the task.

6. I thank Igor Douven for this point.

7. The author thanks Igor Douven for hosting fruitful research stays at his Formal Epistemology Project at the University of Leuven. This work is financially supported by the FWF-project P20209 "Mental probability logic," the DFG grant PF 740/2–1, and the Alexander von Humboldt Foundation.

References

Adams, E.W. 1975. *The Logic of Conditionals*, Dordrecht: Reidel.

———. 2005. "What Is at Stake in the Controversy Over Conditionals," in G. Kern-Isberner, W. Rödder, and F. Kulmann, eds., *Conditionals, Information, and Inference*, Volume 3301 of *LNAI*, Berlin, Heidelberg: Springer, 1–11.

Angell, R.B. 2002. *A-logic*. Lanham: University Press of America.

Bennett, J. 2003. *A Philosophical Guide to Conditionals*, Oxford: Oxford University Press.

Biazzo, V. and A. Gilio 2000. "A Generalization of the Fundamental Theorem of De Finetti for Imprecise Conditional Probability Assessments," *International Journal of Approximate Reasoning* 24 (2–3): 251–72.

Coletti, G. and R. Scozzafava 2002. *Probabilistic Logic in a Coherent Setting*, Dordrecht: Kluwer.

De Finetti, B. 1974 [1970]. *Theory of Probability*, Volume 1, 2, Chichester: John Wiley and Sons.

———. 1980. *Foresight: Its Logical Laws, Its Subjective Sources* (1937), in H.J. Kyburg and H.E. Smokler, eds., *Studies in Subjective Probability*, Huntington, NY: Robert E. Krieger Publishing Company, 55–118.

Douven, I. 2008. "The Evidential Support Theory of Conditionals," *Synthese* 164 (3): 19–44.

Evans, J. St. B.T., S.E. Newstead, and R.M.J. Byrne 1993. *Human Reasoning*: *The Psychology of Deduction*, Hove: Lawrence Erlbaum.

Evans, J. St. B.T. and D.E. Over 2004. *If*. Oxford: Oxford University Press.

Feltz, A. 2009. "Experimental Philosophy," *Analyse & Kritik* 2: 201–19.

Fugard, A.J.B., N. Pfeifer, and B. Mayerhofer 2011. "Probabilistic Theories of Reasoning Need Pragmatics Too: Modulating Relevance in Uncertain Conditionals," *Journal of Pragmatics* 43: 2034–42.

Fugard, A.J.B., N. Pfeifer, B. Mayerhofer, and G.D. Kleiter, 2011, "How People Interpret Conditionals: Shifts Towards the Conditional Event," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37 (3), 635–48.

Grice, H.P. 1975. "Logic and Conversation," in P. Cole and J.L. Morgan, eds., *Syntax and Semantics, Volume 3: Speech Acts*. New York: Seminar Press.

Jackson, F. ed., 1991. *Conditionals*, Oxford: Oxford University Press.

Johnson-Laird, P.N. and R.M.J. Byrne 2002. "Conditionals: A Theory of Meaning, Pragmatics, and Inference," *Psychological Review* 109 (4): 646–78.

Knobe, J. and S. Nichols, eds., 2008. *Experimental Philosophy*, Oxford: Oxford University Press.

Lad, F. 1996. *Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction*, New York: Wiley.

Lewis, D. 1976. "Probabilities Of Conditionals and Conditional Probabilities," *Philosophical Review* 85: 297–315. Reprint with postscript in Jackson (1991, 76–101); the page references are to the reprint.

Lukasiewicz, J. 1957. *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*, 2nd ed., Oxford: Oxford University Press.

McCall, S. 1966. "Connexive Implication," *Journal of Symbolic Logic* 31: 415–33.

Oaksford, M. and N. Chater 2009. Précis of "Bayesian rationality: The Probabilistic Approach to Human Reasoning," *Behavioral and Brain Sciences* 32, 69–84.

Oberauer, K. 2006. "Reasoning with Conditionals: A Test of Formal Models of Four Theories," *Cognitive Psychology* 53: 238–83.

Over, D.E., C. Hadjichristidis, J. St. B. T. Evans, S.J. Handley, and S. Sloman 2007. "The Probability of Causal Conditionals," *Cognitive Psychology* 54: 62–97.

Pfeifer, N. and G.D. Kleiter 2003. "Nonmonotonicity and Human Probabilistic Reasoning," in *Proceedings of the Sixth Workshop on Uncertainty Processing*, Hejnice, September 24–27, 221–34.

———. 2005a. "Coherence and Nonmonotonicity in Human Reasoning," *Synthese* 146(1–2): 93–109.

———. 2005b. "Towards a Mental Probability Logic," *Psychologica Belgica* 45(1): 71–99.

———. 2006a. "Inference in Conditional Probability Logic," *Kybernetika* 42, 391–404.

———. 2006b. "Is Human Reasoning About Nonmonotonic Conditionals Probabilistically Coherent?" in *Proceedings of the Seventh Workshop on Uncertainty Processing*, Mikulov, September 16–20, 2006, 138–50.

———. 2007. "Human Reasoning with Imprecise Probabilities: Modus Ponens and Denying the Antecedent," in G. De Cooman, J. Vejnarová, and M. Zaffalon, eds., *Fifth International Symposium on Imprecise Probability: Theories and Applications*, Prague, Czech Republic, 347–56.

———. 2009. "Framing Human Inference by Coherence Based Probability Logic," *Journal of Applied Logic* 7(2): 206–17.

———. 2010. "The Conditional in Mental Probability Logic," in M. Oaksford and N. Chater, eds., *Cognition and Conditionals: Probability and Logic in Human Thought*, Oxford: Oxford University Press,153–73.

———. 2011. "Uncertain Deductive Reasoning," in K. Manktelow, D.E. Over, and S. Elqayam, eds., *The Science of Reason: A Festschrift for Jonathan St B.T. Evans*, Hove: Psychology Press, 145–66.

Phillips, J. 2011. "The Experimental Philosophy Webpage," http://experimentalphilosophy.org/ ExperimentalPhilosophy.html. retrieved April 2011.

Ramsey, F.P. 1978. "Truth and probability (1926)," in D.H. Mellor, ed., *Foundations. Essays in Philosophy, Logic, Mathematics and Economics*, London: Routledge & Kegan Paul, 58–100.

Walley, P. 1991. *Statistical Reasoning with Imprecise Probabilities*, London: Chapman and Hall.

Wansing, H. 2010. "Connexive Logic," in E.N. Zalta, ed., *The Stanford Encyclopedia of Philosophy* (Fall 2010 ed.).

Uncertain deductive reasoning

Pfeifer, N. & Kleiter, G. D.
University of Salzburg, Department of Psychology
niki.pfeifer@sbg.ac.at gernot.kleiter@sbg.ac.at

Abstract

Probabilistic models have started to replace classical logic as the standard reference paradigm in human deductive reasoning. Mental probability logic emphasizes general principles where human reasoning deviates from classical logic, but agrees with a probabilistic approach (like nonmonotonicity or the conditional event interpretation of conditionals).

This contribution consists of two parts. In the first part we discuss general features of reasoning systems including consequence relations, how uncertainty may enter argument forms, probability intervals, and probabilistic informativeness. These concepts are of central importance for the psychological task analysis. In the second part we report new experimental data on the paradoxes of the material conditional, the probabilistic MODUS PONENS, the COMPLEMENT task, and data on the probabilistic truth table task. The results of the experiments provide evidence for the hypothesis that people represent indicative conditionals by conditional probability assertions.

1 Introduction

The title of this contribution appears paradoxical. Deduction is truth preserving and not uncertain. The alleged paradox is clarified in the following sections.

A descriptive theory of human reasoning needs a normative theory in the background. The normative background provides rationality norms and specifies the correct answer to a reasoning task relative to these norms. It serves to provide the language and the inference rules to solve the problem. Moreover, it stimulates psychological hypotheses and guides the construction of experimental tasks. The shift from classical logic to alternative normative systems leads to new experimental paradigms. The choice of an appropriate normative framework is a fundamental problem and needs careful consideration before running psychological experiments. If the normative framework is ignored in the process of psychological model building, then one runs into the danger of not knowing *what* one investigates: in the worst case, the resulting experimental data become uninterpretable.

The research on human reasoning is a process that proceeds by asking questions and setting goals, by designing experiments and selecting experimental tasks, and by interpreting the results within a theoretical framework. All these steps are related to a normative paradigm. For many years it was taken for granted that classical logic is the appropriate normative paradigm for research on human reasoning. Modern logic and mathematical computer science, however, have developed many systems to describe reasoning tasks, properties of knowledge representation, and inference. Today there are many systems of "rational" reasoning. It is reasonable to consider some of these logical systems while asking questions and setting goals, while designing experiments and selecting experimental tasks, and while interpreting the results within a theoretical framework. Probability logics provide one class of such systems. The present contribution argues that the *coherence based probability logic* is a fruitful normative frame for investigating human reasoning. Before we illustrate the application of probability logic to psychology, we discuss some theoretical notions that are important for human reasoning research.

Where do classical logic and probability logic differ and where not? First, what is common to both.

- Propositions (or events) are true or false. They have Boolean truth values.

	State of the world		*Material conditional*	*Conjunction*	*Conditional event*
	A	B	$A \supset B$	$A \wedge B$	$B\|A$
s_1	*true*	*true*	*true*	*true*	*true*
s_2	*true*	*false*	*false*	*false*	*false*
s_3	*false*	*true*	*true*	*false*	undetermined
s_4	*false*	*false*	*true*	*false*	undetermined

Table 1: Truth tables of three prominent psychological interpretations of the indicative "if—, then—". Only rows s_3 and s_4 distinguish between the three interpretations.

- Inferences from premises to conclusions are deductive. Probability logic investigates argument forms like MODUS PONENS and MODUS TOLLENS, or various kinds of syllogisms in an analogous way as in classical logic.

Major differences concern the following points:

- In probability logic propositions are assigned degrees of belief. The degrees of belief are subjective probabilities. Subjective probabilities are coherent descriptions of partial knowledge states.
- Probability assessments should be coherent. A probability assessment is coherent if it cannot lead to bets with sure loss. There are different kinds of coherence.
- Probabilistic inference transmits probabilities from the premises to the conclusions. The inference itself is deductive and not probabilistic. It follows the rules of mathematics, like solving systems of linear equations or linear programming.
- There are a number of rather general properties in which both systems differ. Classical logic is monotone, probability logic is defeasible and nonmonotone. Classical logic is truth functional, a property that does not apply to probability logic. Systems witch such meta-properties are investigated in artificial intelligence to mimic human reasoning. They are of special interest for psychological modeling.

During the last years several probabilistic models of human reasoning were proposed (Oaksford & Chater, 1998, 2007a; Evans & Over, 2004). These models do of course have many features in common with the present approach. One common feature is that the probabilities of conditionals, *P(If A, then B)*, are interpreted as conditional probabilities, $P(B|A)$, as opposed to the probability of material conditionals, $P(A \supset B)$. Table 1 presents three prominent psychological interpretations of the indicative "if—, then—". The interpretation as a conditional event received compelling empirical support (e.g., Liu, Lo, & Wu, 1996; Evans, Handley, & Over, 2003; Over & Evans, 2003; Oberauer & Wilhelm, 2003). The dominant experimental paradigm for investigating the understanding of uncertain indicative conditionals is the probabilistic truth table task. We proposed an alternative experimental paradigm by studying probabilistic argument forms. Our own approach is called "mental probability logic" (Pfeifer & Kleiter, 2005a, 2005b, in pressb, in pressa) and differs from other probabilistic approaches with respect to the following points:

- Our commitment to subjective probability and coherence is in the tradition of de Finetti (Coletti & Scozzafava, 2002). In combination with recent developments in computer science (like description logics (Baader, Calvanese, McGuinness, Nardi, & Patel-Schneider, 2007)), the coherence approach leads to powerful systems of knowledge representation and inference (Gilio, 2002; Biazzo, Gilio, Lukasiewicz, & Sanfilippo, 2005; Lukasiewicz, 2005, 2008).
- In our approach it often occurs that the conclusions of an argument obtain lower and upper probabilities and not point probabilities. We investigate interval probabilities.
- In our approach the consequence relation itself is deductive and not probabilistic.

The present contribution consists of two parts. In the first part we discuss a selection of logical features of reasoning systems that are of central importance for the psychological task analysis.

In the second part we report experimental data on the paradoxes of the material conditional, the probabilistic MODUS PONENS, the COMPLEMENT task, and data on the probabilistic truth table task.

2 General features of reasoning systems

Human thinking and reasoning may lead to outstanding and fascinating achievements. Take as an example the proof of Fermat's Last Theorem. It took more than 350 years to come up with a proof (Wiles, 1995) of the famous conjecture that Euler or Poincaré could not solve. We think that any reasonable theory of human reasoning should strive to explain how we solve difficult problems. Solving difficult problems requires a minimum of systematicity. Considering various normative background systems one should be careful to chose a system that does not exclude smart solutions right from the beginning. While evaluating formal systems for their appropriateness for psychological modeling, one should look out for "competent" and rich systems. Description logic is one such family of systems (Baader et al., 2007). It has replaced semantic networks for knowledge representation, it has many extensions (probabilistic, possibilistic, autoepistemic, etc.); so it is a rich system. There is software for real-world applications to build ontologies.[1] In our view it is important to look at principles developed in such fields to learn more also about features and constraints of human reasoning.

Features that cannot be expressed in classical logic include procedural mechanisms, defaults, closed world reasoning, assumptions about missing information, non-truth functionality, non-monotonicity, defeasible reasoning (in a wide sense), uncertainty, probability, vagueness, and fuzziness.

Several disciplines investigate knowledge representation and reasoning. Psychology is only one of them. During the last fifty years logicians and computer scientists have developed a large repertoire of logics and formal systems of reasoning (Van Harmelen, Lifschitz, & Porter, 2008), recently especially in connection with the semantic web projects. Usually these systems are "non-classical", that is they are extensions of classical logic or they explicitly violate principles of classical logic. A typical example is nonmonotonic reasoning.

During the last few years these non-classical systems have had a strong influence on which questions and experimental tasks are investigated. Extensions of classical logic include causal reasoning, autoepistemic reasoning, epistemic logic, dynamic logic, deontic logic, temporal logic, fuzzy logic, possibilistic logic, and—last but not least—probability logic.

3 Deduction and Uncertainty

One of the basic concepts of classical logic is validity ("cl-validity" for short). Cl-validity guarantees that the truth of the premises propagates to the conclusion. As an example of an argument that is *valid* according to classical logic, consider the following instance of the MODUS PONENS:

$\mathfrak{P}1$	**If** shape X is a triangle, **then** shape X is blue.	
$\mathfrak{P}2$	Shape X is a triangle.	
	———————————————	cl-valid
$\mathfrak{C}$	Shape X is blue.	

$\mathfrak{P}1$ and $\mathfrak{P}2$ denote the premises and $\mathfrak{C}$ denotes the conclusion. The horizontal line denotes the consequence relation. If the conditional in $\mathfrak{P}1$ (**If** A, **then** B) is formalized as a material conditional, $A \supset B$, then this argument is cl-valid, since $\mathfrak{C}$ is true under *all* interpretations that assign *true* to both $\mathfrak{P}1$ and $\mathfrak{P}2$.

Cl-validity is a meta-property of arguments, and not of the conclusion of arguments. However, if all premises are true *and* if an argument is cl-valid, then the conclusion must be true. If at least one premise is false, then classical logic *does not* tell us whether the conclusion is true or false. This distinction between the two questions—(i) is an argument cl-valid?, and (ii) is the conclusion true if all premises are true?—leads to two classes of experimental tasks.

[1] For example `protégé`, see `http://protege.stanford.edu/`.

	MODUS PONENS		DENYING THE ANTECEDENT	
	affirmative	*negated*	*affirmative*	*negated*
$\mathfrak{P}1$:	$A \supset B$	$A \supset B$	$A \supset B$	$A \supset B$
$\mathfrak{P}2$:	A	A	$\neg A$	$\neg A$
$\mathfrak{C}$	B	$\neg B$	$\neg B$	B
cl-valid:	yes	no	no	no
$V(\mathfrak{C})$	*true*	*false*	undetermined	undetermined

Table 2: MODUS PONENS, DENYING THE ANTECEDENT, and their respective affirmative and negated versions. $\mathfrak{P}1$ and $\mathfrak{P}2$ denote the premises and $\mathfrak{C}$ denotes the conclusion. A and B denote propositions. $\supset$ and $\neg$ denote the material conditional and negation, respectively, and are defined as usual. *cl-valid* denotes classical logical validity. V denotes the logical valuation-function V under all interpretations that assign *true* to each premise. If the antecedent, A, of the conditional premise is false, then the truth value of the conclusion is not determined. (Adapted from Pfeifer & Kleiter, 2007, p. 347.)

The first class of experimental tasks requires the judgment of cl-validity. These tasks require a choice between two options: the *argument* is (i) cl-valid or (ii) not cl-valid. Cl-validity is a meta-property of the (form of the) whole argument.

The second class of experimental tasks requires a judgment concerning the conclusion of an argument. In such tasks the participants are instructed to assume that each premise is true and to decide whether the *conclusion* is (i) true, (ii) false, or (iii) whether one cannot decide whether the conclusion is true or false. This task requires a choice between three options. Option (iii) is of central importance but unfortunately not always made available for the participants (especially in the "possibilities" form of the truth table tasks, which is often used by proponents of the mental model theory). Normatively, if the argument is not cl-valid, then the truth value of the conclusion is undetermined (for example DENYING THE ANTECEDENT) or necessarily false (if, for example, the conclusion of MODUS PONENS is negated). The normative relationships between both classes of tasks are explained in Table 2.

Cl-validity is a complex and abstract concept. It is a *meta-property* of arguments. To judge whether an argument is cl-valid is a rather artificial task which is hardly ever performed in everyday life. People usually do not evaluate abstract meta-properties of arguments. Rather, people focus on the conclusion. This does not mean that we think that people are always prone to *belief biases*[2]. Rather, people are typically concerned with the problem of evaluating a concrete conclusion in the light of the given evidence. People try to infer the conclusion deductively from the premises. Moreover we assume that everyday life arguments are almost always uncertain and defeasible. Conclusions are often retracted in the light of new evidence, and uncertainty is at least implicitly present in almost all common sense arguments. How does this uncertainty enter the normative models of everyday life arguments?

We discuss two different ways of introducing uncertainty to argument forms. The first option is to introduce an uncertain consequence relation between the premises and the conclusion. The second option is to attach probabilities to the premises and to keep the consequence relation deductive. Oaksford, Chater, and Larkin (2000) opted for the first option in their analysis of the four conditional syllogisms (MODUS PONENS, MODUS TOLLENS, AFFIRMING THE CONSEQUENT, and DENYING THE ANTECEDENT). They represent a conditional syllogism by the conditional probability of the conclusion given the categorical premise, $P(\text{conclusion} \mid \text{categorical premise})$. This makes the consequence relation uncertain. As an example consider the MODUS PONENS,

$\mathfrak{P}1$ **If** shape X is a triangle, **then** shape X is blue.
$\mathfrak{P}2$ Shape X is a triangle.
.... uncertain
$\mathfrak{C}$ Shape X is blue.

[2] "Belief bias" denotes the tendency to evaluate an argument just by the believability of the conclusion and to ignore the premises (Evans, Barston, & Pollard, 1983)).

Oaksford et al. (2000) interpret this argument as the conditional probability of $\mathfrak{C}$ given $\mathfrak{P}2$, P(Shape X is blue. | Shape X is a triangle.). The conditional premise ($\mathfrak{P}1$) does not enter the model. This analysis does not preserve the original propositional structures of the arguments. Liu (2003) and Oaksford and Chater (2007b) modified this model such that the conclusion is conditionalized on the conditional ($\mathfrak{P}1$) and on the categorical premise ($\mathfrak{P}2$).

The normative treatment of the uncertain consequence relation beyond the four popular conditional syllogisms, however, is challenging. Adam's p-validity (Adams, 1975) is an example of *qualitatively* uncertain consequence relations. Systems of nonmonotonic reasoning provide other examples (Antoniou, 1997). The definition of satisfactory consequence relations that are *quantitatively* uncertain is an open problem.

We next turn to our approach. Because of its normative frame (probability logic based on coherence) we call it "mental probability logic" (Pfeifer & Kleiter, 2005b, in pressb, in pressa). In our approach the consequence relation is deductive and the conclusion is uncertain. As an example consider the probabilistic MODUS PONENS:

$$\begin{array}{ll} \mathfrak{P}1 & P(\text{Shape X is blue.} \mid \text{Shape X is a triangle.}) = x \\ \mathfrak{P}2 & P(\text{Shape X is a triangle.}) = y \\ \hline \mathfrak{C} & xy \quad \leq \quad P(\text{Shape X is blue.}) \quad \leq \quad xy + 1 - y \end{array} \quad \text{deductively valid}$$

The argument form contains two binary variables. A complete probabilistic assessment requires three probabilities, say x, y, and z of three logically independent events. The premises of the probabilistic MODUS PONENS specify only two probability values, x and y. The coherent probability of the conclusion is imprecise. It is constrained by a lower and an upper bound. In $\mathfrak{P}1$ the conditional is interpreted as a conditional event, "Shape X is blue. | Shape X is a triangle.". If the conditional in $\mathfrak{P}1$ is interpreted as a material conditional, "Shape X is a triangle. $\supset$ Shape X is blue.", then the probability propagation rule is different:

$$\begin{array}{ll} \mathfrak{P}1' & P(\text{Shape X is a triangle.} \supset \text{Shape X is blue.}) = x \\ \mathfrak{P}2 & P(\text{Shape X is a triangle.}) = y \\ \hline \mathfrak{C} & \max\{0, x + y - 1\} \quad \leq \quad P(\text{Shape X is blue.}) \quad \leq \quad x \end{array} \quad \text{deductively valid}$$

The difference between the material conditional and the conditional event is that the truth value of the conditional event is undetermined if the antecedent of the conditional is false (see Table 1). This affects, of course, the probability propagation rules. Probability logic provides the language to represent and reason from different interpretations of the premises.

How people interpret conditionals is debated in the literature. We discuss two prominent positions. The first position favors the material conditional interpretation. The second position favors the conditional event interpretation. Mental model theory takes the first position (Johnson-Laird & Byrne, 2002). The theory postulates that people represent indicative conditionals either as implicit mental models or as explicit mental models. To be precise, this is postulated only for those conditionals that are as independent as possible from context or background knowledge. According to the mental model theory, the truth conditions of the implicit mental model coincide with the truth conditions of the conjunction, $\wedge$. The truth conditions of the explicit mental model coincide with the truth conditions of the material conditional, $\supset$. For the truth conditions see Table 1. Explicit mental models have a higher working memory demand and are harder to process than implicit mental models. Therefore, the theory predicts that people usually form more implicit mental models than explicit ones.

Barrouillet and Lecas (1999) instructed participants to list possible truth-table cases that are consistent with an indicative "if—, then—". The results are consistent with the material conditional interpretation, and are treated as strong evidence for the mental model theory. However, these results are *ambiguous* since they are consistent with the conditional event interpretation as well. Normatively, both interpretations differ only in the truth table cases s_3 and s_4, where the material conditional is *true* but the conditional event is *undetermined* (see Table 1). The truth table cases that are "possible" are consistent with both the truth table cases that are *true* and those that are *undetermined*. Therefore, this version of a truth table task cannot differentiate among the material conditional and the conditional event interpretation of the indicative "if—, then—".

The second position postulates that people interpret indicative conditionals as a conditional event. A specialty of our approach is, that the conditional event, $B|A$, is *basic* in mental probability logic. Probabilities are assigned *directly* to the conditional event. This is psychologically plausible and reduces the working memory demand: one does not need to process the joint ($P(A \wedge B)$) and the marginal ($P(A)$) probabilities.

Many empirical studies support the conditional event interpretation of conditionals (Evans et al., 2003; Over & Evans, 2003; Oberauer & Wilhelm, 2003). These studies use an experimental paradigm that investigates *complete* probabilistic knowledge. Probabilistic truth table tasks present (or ask for) the probabilities of all truth table cases. One version of the probabilistic truth table task presents the probabilities of $A \wedge B$, $A \wedge \neg B$, $\neg A \wedge B$, and of $\neg A \wedge \neg B$ to the participants. The participants infer the probability of a conditional. Since complete probabilistic knowledge is provided, the probabilities of conjunctions, material conditionals, and conditional events are *point values*. Another version of the probabilistic truth table task provides the point probability of a conditional and the participants rate the probabilities of all four truth table cases. Again, this results in a task that consists of inferences about complete probabilistic knowledge.

Our approach uses an alternative experimental paradigm. We investigate selected argument forms and present *incomplete* probabilistic knowledge in the premises. In our MODUS PONENS tasks, for example, we present only two probabilities to the participants: $P(B|A)$ and $P(A)$. As explained above, the probability of the conclusion, $P(B)$, is an *interval probability* and not a point probability. If $P(B|\neg A)$ were also given, then the probability of the conclusion would be a point value. $P(B|\neg A)$, however, would be an additional premise and the resulting task would not map the MODUS PONENS any more. Therefore, we investigate incomplete probabilistic knowledge. We observed good agreement between the responses and the coherent probability intervals of the probabilistic MODUS PONENS (Pfeifer & Kleiter, 2007, in pressb). How the participants understand conditionals can be concluded from the inferences they draw.

An important problem of probabilistic argument forms is whether the coherent probability of the conclusion is constrained by the premises. If the coherent probability of the conclusion of an argument form is not necessarily equal to the unit interval, $[0, 1]$, then we call this argument form "*probabilistically informative*". We call an argument form "*probabilistically non-informative*" if the assignment of the unit interval to its conclusion is coherent for all probability assessments of the premises (Pfeifer & Kleiter, in pressb). If an argument is probabilistically non-informative, then one cannot infer anything about the probability of the conclusion (except that it is between zero and one). Whether an argument is probabilistically informative can depend upon how the conditional is interpreted. We discuss this fact by two of the so-called "paradoxes of the material conditional":

PARADOX 1: *from* $\boxed{B}$ *infer* $\boxed{A \supset B}$.
PARADOX 2: *from* $\boxed{\neg A}$ *infer* $\boxed{A \supset B}$.

Both PARADOX 1 and PARADOX 2 are cl-valid. There is nothing paradoxical about this. However, the paradoxes arise if indicative natural language conditionals are interpreted as material conditionals. Consider, for example, the following instances:

Instance of PARADOX 1:
From $\boxed{\textit{It is raining.}}$ *infer* $\boxed{\textit{If I'm happy, then it is raining.}}$.

Instance of PARADOX 2:
From $\boxed{\textit{I'm not happy.}}$ *infer* $\boxed{\textit{If I'm happy, then it is raining.}}$.

Such examples are well known in the history of logic (Lewis, 1918). Obviously both inferences are odd, but cl-valid if the conclusion is interpreted as a material conditional. Table 3 presents the paradoxes, and three probability logical interpretations of the "if—, then—". The probability of the conclusion is constrained under the material conditional interpretation and under the conjunction interpretation. Under the conditional event interpretation, however, the unit interval, $[0, 1]$, is coherent for all probability assessments of the premise. Thus, the paradoxes are probabilistically non-informative if the conditional is interpreted as a conditional event. If the

		Premise		Conclusion			Probabilistically informative
P1:	(a)	B	$\models$		$A \supset B$		
	(b)	$P(B) = x$	$\models$	$x \leq$	$P(A \supset B)$	≤ 1	yes
	(c)	$P(B) = x$	$\models$	$0 \leq$	$P(A \wedge B)$	$\leq x$	yes
	(d)	$P(B) = x$	$\models$	$0 \leq$	$P(B\|A)$	≤ 1	no
P2:	(a)	$\neg A$	$\models$		$A \supset B$		
	(b)	$P(\neg A) = x$	$\models$	$x \leq$	$P(A \supset B)$	≤ 1	yes
	(c)	$P(\neg A) = x$	$\models$	$0 \leq$	$P(A \wedge B)$	$\leq 1 - x$	yes
	(d)	$P(\neg A) = x$	$\models$	$0 \leq$	$P(B\|A)$	≤ 1	no
NP1:	(a)	B	$\not\models$		$A \supset \neg B$		
	(b)	$P(B) = x$	$\models$	$1 - x \leq$	$P(A \supset \neg B)$	≤ 1	yes
	(c)	$P(B) = x$	$\models$	$0 \leq$	$P(A \wedge \neg B)$	$\leq 1 - x$	yes
	(d)	$P(B) = x$	$\models$	$0 \leq$	$P(\neg B\|A)$	≤ 1	no
NP2:	(a)	$\neg A$	$\models$		$A \supset \neg B$		
	(b)	$P(\neg A) = x$	$\models$	$x \leq$	$P(A \supset \neg B)$	≤ 1	yes
	(c)	$P(\neg A) = x$	$\models$	$0 \leq$	$P(A \wedge \neg B)$	$\leq 1 - x$	yes
	(d)	$P(\neg A) = x$	$\models$	$0 \leq$	$P(\neg B\|A)$	≤ 1	no

Table 3: Two paradoxes of the material conditional (P1 and P2) and their respective negated versions (NP1 and NP2). For each paradox four interpretations are given: (a) classical logic, and the three probability logical interpretations of the "*if—, then—*" (b) as a material conditional, (c) as a conjunction, and (d) as a conditional event. "$\models$" denotes deductive validity, "$\not\models$" denotes "not deductively valid". Only under the conditional event interpretation (d) the paradoxes are probabilistically non-informative, and therefore *not* paradoxical.

conditional is interpreted as a material conditional or as a conjunction, then the paradoxes are probabilistically informative (Pfeifer & Kleiter, 2006). Under the conditional event interpretation, however, the paradoxical nature of these argument forms disappears.

If the antecedent of the conditional in the respective conclusion is negated, we call the resulting argument forms the "negated versions" of the paradoxes (see Table 3). Both paradoxes and their negated versions provide different probability logical predictions on the three popular interpretations of the natural language "if—, then—". According to the mental model theory most participants will make inferences that correspond to the predictions of the conjunction interpretation (implicit mental model) or to the material conditional interpretation (explicit mental model) of the "if—, then—" (Johnson-Laird & Byrne, 2002). Mental probability logic predicts that most participants make inferences that correspond to the predictions of the conditional event interpretation of the "if—, then—". We investigate these predictions in the following sections.

4 Experiment 1: Two paradoxes of the material conditional

4.1 Introduction and Method

Experiment 1 investigates PARADOX 1 and PARADOX 2, and their respective negated versions, see Table 3. Between each paradox task we presented MODUS PONENS tasks. The respective negated versions of these argument forms were presented in the second half of the experiment.

4.1.1 Participants

Thirty-two students of the University of Salzburg were payed five € each for their participation in the experiment (mean age: 23 years, 21 female, 11 male participants). Psychology students, mathematics students and students with a background in logic were not included in the sample. To ensure an atmosphere for thinking and reasoning, each participant was tested individually in an experimental room in the department of psychology. The participants were instructed to take their time and to think carefully about each problem.

4.1.2 Material

PARADOX 1 (*From* $\boxed{B}$ *infer* $\boxed{\textit{If A, then B}}$) and PARADOX 2 (*From* $\boxed{\neg A}$ *infer* $\boxed{\textit{If A, then B}}$; see Table 3) were translated into cover-stories. As an example consider the following problem, which contains an instance of PARADOX 1:

> Simon works in a factory that produces playing cards. He is responsible for what is printed on the cards.
>
> On each card, there is a *shape* (triangle, square, . . .) of a certain *color* (green, blue, . . .), like:
>
> - green triangle, green square, green circle, . . .
> - blue triangle, blue square, . . .
> - red triangle, . . .
>
> Imagine that a card got stuck in the printing machine. Simon cannot see what is printed on this card. Since Simon observed the card production during the whole day, he is
>
> $\boxed{\text{A}}$ 90% certain: There is a *square* on this card.
>
> *Considering* $\boxed{\text{A}}$*, how certain can Simon be that the following sentence is true?*
>
> $\boxed{\textbf{If}\text{ there is a }\textit{red}\text{ shape on this card, }\textbf{then}\text{ there is a }\textit{square}\text{ on this card.}}$

The premise is to the right of " $\boxed{\text{A}}$ " and the box contains the conditional in the conclusion. We were careful to ensure that the conditionals are as independent as possible from background knowledge. According to the mental model theory, the truth conditions of such conditionals coincide either with those of the conjunction or with those of the material conditional (Johnson-Laird & Byrne, 2002).

First the participants decided whether the inference is probabilistically informative or not. If they judged it not to be, then they continued to the next problem. If the argument was identified as probabilistically informative, then the participants rated their degree of belief in the conclusion. This two-step procedure was phrased as follows:

> Considering $\boxed{\text{A}}$, can Simon infer—at all—how certain he can be that the sentence in the box is true?
>
> ☐ NO, Simon cannot infer his certainty, since anything between 0% and 100% is possible.
>
> ☐ YES, Simon can infer his certainty.
>
> *In case you ticked* YES, *please fill in*
>
> Simon can be certain *from* at least ___% *to* at most ___% that the sentence in the box is true.
>
> 0 ———— 100 %

The thirty-two participants were randomly assigned to two conditions. In the first condition we presented five instances of PARADOX 1 ($n_1 = 16$) and in the second condition we presented five instances of PARADOX 2 ($n_2 = 15$, one participant was excluded from the data analysis because of misunderstandings of the instructions). The versions differed in the uncertainties attached to the premises. The uncertainties were: "60%", "70%", "90%", "pretty sure" (German original: ziemlich sicher), and "absolutely certain" (absolut sicher). Between each paradox we presented a MODUS PONENS task. The uncertainty of the categorical premise ($\boxed{P(A)}$) of the MODUS PONENS tasks was kept constant at 100%. The uncertainties of the conditional premise ($\boxed{P(\textit{if } A, \textit{ then } B)}$) were: "90%", "70%", "80%", "pretty sure", and "absolutely certain". After the ten tasks just described, we presented the same tasks in the same order again, with the difference that the B term in the conclusions was negated. In the negated MODUS PONENS tasks the conclusion had the form $\neg B$ (rather than B), and in both versions of the negated paradoxes the conditional had the form *If A, then* $\neg B$ (rather than *If A, then B*).

4.2 Results

Averaging across subjects and items, more than 80% of the responses identified the MODUS PONENS correctly as probabilistically informative (see Table 4). Of all those participants who inferred correctly that the MODUS PONENS is probabilistically informative, all participants—except one—inferred correctly that it is "absolutely certain"/"pretty sure" that the conclusion is true. In the negated MODUS PONENS tasks, all participants—except two—inferred correctly that it is "absolutely certain"/"pretty sure" that the conclusion is not true. In these tasks most participants understand that complementary probabilities should add up to one. The good performance in our MODUS PONENS tasks corresponds to the high endorsement rates of the non-probabilistic MODUS PONENS tasks reported in the literature (Evans, Newstead, & Byrne, 1993).

The majority of the participants (more than 70% on the average) inferred that all versions of the paradoxes are not probabilistically informative (see Table 4). This speaks for the conditional event interpretation of the conditional. According to the mental model theory, most participants should make probabilistically informative inferences. Their inferences should conform to the conjunction (implicit mental model) or to the material conditional interpretation (explicit mental model). These predictions are not confirmed by the data.

To our knowledge, this is the first experiment on these two paradoxes of the material conditional. The high percentage of probabilistically non-informative responses is remarkable from a pragmatic point of view: they are obviously not influenced by conversational implicatures that would call for informative inferences. Rather, the participants seem to interpret the conditional as a conditional event and therefore understand that one cannot infer anything.

In another experiment we investigated the following argument form, which is called "monotonicity" or "premise strengthening": *From* $\boxed{\textit{If } A, \textit{ then } B}$ *infer* $\boxed{\textit{If } A \textit{ and } C, \textit{ then } B}$. This argument form is another well known paradox of the material conditional. The results are in line with the present data: we observed wide and practically non-informative interval responses (Pfeifer & Kleiter, 2003).

5 Experiment 2

5.1 Introduction and method

In the following sections we report selected parts of a study which consisted of thirty seven different tasks. One main purpose of Experiment 2 was to attempt to replicate the results on the paradoxes of Experiment 1. We report these tasks and focus on other

					Affirmative argument forms					
	P_{60}	P_{70}	P_{90}	P_{ps}	P_{ac}	MP_{90}	MP_{70}	MP_{80}	MP_{ps}	MP_{ac}
P1	63	81	69	69	69	63	88	81	75	94
P2	73	73	73	80	67	73	73	87	80	93
					Negated argument forms					
	P_{60}	P_{70}	P_{90}	P_{ps}	P_{ac}	MP_{90}	MP_{70}	MP_{80}	MP_{ps}	MP_{ac}
P1	75	69	63	75	44	81	88	88	69	88
P2	87	87	87	67	67	80	87	73	93	93

Table 4: Percentages of probabilistic non-informativeness responses in the Paradox (P) tasks, and of probabilistic informativeness responses in the Modus Ponens (MP) tasks. P1: PARADOX 1 ($n_1 = 16$), P2: PARADOX 2 ($n_2 = 15$). The indices denote the percentages presented in the premises, "ac" denotes "absolutely certain", "ps" denotes "pretty sure". The indices of the MODUS PONENS tasks denote the percentages presented in the conditional premise.

Response	First P1	Second P1	First NP1	Second NP1
non-informative	13	15	13	12
conclusion is true	3	2	0	1
conclusion is false	3	2	6	6

Table 5: Response frequencies of the two affirmative PARADOX 1 tasks (P1) and the two negated PARADOX 1 tasks (NP1) of Experiment 2 ($n_3 = 19$).

tasks that investigate the representation of conditionals as well. Moreover, we present new data on the probabilistic truth table task.

5.1.1 Participants

Forty students of the University of Salzburg were payed five € each. They were divided into two groups of twenty participants each. Most tasks of group I investigated affirmative arguments and most tasks of group II included negations. The participants were tested individually in an experimental room. One participant was not included in the data analysis because he misunderstood the instructions. As in Experiment 1, the participants were instructed to take their time and to think carefully about each problem.

5.2 Material and results

The two paradoxes In Experiment 2 we investigated the two paradoxes of the material conditional (see Table 3). The thematic content of all tasks and the response modalities were the same as in Experiment 1. The uncertainty in each premise was formulated verbally as "pretty sure" (ziemlich sicher).

The nineteen participants of group I solved PARADOX 1 twice and the negated version of PARADOX 1 twice (see Table 3). Between each paradox task, we presented tasks that investigate different argument forms. Table 5 presents the response frequencies of the affirmative and of the negated PARADOX 1 tasks.

The majority (63%-79%) of the participants understand that PARADOX 1 and the negated version of PARADOX 1 are probabilistically non-informative. This speaks for the conditional probability interpretation of the "if—, then—". The data replicate the findings of Experiment 1.

Response	First P2	Second P2	First NP2	Second NP2
non-informative	15	18	18	18
conclusion is true	1	1	0	1
conclusion is false	4	1	2	1

Table 6: Response frequencies of the two affirmative PARADOX 2 tasks (P1) and the two negated PARADOX 2 tasks (NP2) of Experiment 2 ($n_4 = 20$).

The participants of group II ($n_4 = 20$) solved PARADOX 2 twice and the negated version of PARADOX 2 twice (see Table 3). Between each paradox task, we presented again other tasks. Table 6 presents the response frequencies of the affirmative and of the negated PARADOX 2 tasks.

Almost all participants (75%-90%) understand that PARADOX 2 and the negated version of PARADOX 2 are probabilistically non-informative. This speaks for the conditional probability interpretation of the "if—, then—". The data replicate the findings of Experiment 1.

The complement tasks The paradoxes are probabilistically non-informative, if the conditional is interpreted as a conditional event. They are probabilistically informative under the material conditional interpretation. There are argument forms in which the rôle of the probabilistic informativeness of the different interpretations is interchanged. In the following argument form, the conditional event interpretation is probabilistically informative but the material conditional interpretation is practically non-informative (a wide interval between a value close to zero and one).

$$\textit{From } \boxed{\textit{If } A, \textit{ then } B} \textit{ infer } \boxed{\textit{If } A, \textit{ then } \neg B} \; .$$

Lets call this argument form "COMPLEMENT". If the uncertain "if—then—" of the COMPLEMENT is interpreted as a conditional event, then the resulting argument form is probabilistically informative:

$$\textit{From } \boxed{P(B|A) = x} \textit{ infer } \boxed{(\neg B|A) = 1 - x} \; .$$

If, however, the conditional is interpreted as a material conditional, then the COMPLEMENT is formalized as:

$$\textit{From } \boxed{P(A \supset B) = x} \textit{ infer } \boxed{1 - x \leq P(A \supset \neg B) \leq 1} \; .$$

Thus, if the conditional is interpreted as a material conditional and if the probability of the premise is sufficiently high, then the conclusion is practically non-informative, since the probability of the conclusion is anywhere between a very low value and one.[3]

If the conditional of the COMPLEMENT is interpreted as a conjunction, then it is formalized as:

$$\textit{From } \boxed{P(A \wedge B) = x} \textit{ infer } \boxed{0 \leq P(A \wedge \neg B) \leq 1 - x} \; .$$

If the conditional is interpreted as a conjunction (implicit mental model) and if the probability of the premise is sufficiently high, then the probability of the conclusion is close to zero.

[3]This argument form is strictly speaking probabilistically informative, since it is not the case that for all probability assignments of the premise, the coherent probability of the conclusion is necessarily between zero and one. Therefore we say that this argument form is "practically non-informative".

The predictions concerning the COMPLEMENT task are straightforward. If the participants interpret the conditional as a material conditional, they should give a probabilistically non-informative response. The presented COMPLEMENT task does not distinguish between the conditional event interpretation and the conjunction interpretation. (If the probability of the premise would have been lower, then it could differentiate among all three interpretations.)

In Experiment 2, we presented the COMPLEMENT as the first task to all participants ($n_3 + n_4 = 39$). Twenty-eight participants responded by ticking the box “pretty sure” that the conclusion is false. This corresponds to the conditional event interpretation and to the conjunction interpretation. Nine participants opted for the probabilistic non-informativeness, which corresponds to the material conditional interpretation.

We included also a task that investigates the negated COMPLEMENT. The negated version of the COMPLEMENT corresponds to the trivial inference: *from* $\boxed{\textit{if } A, \textit{ then } B}$ *infer* $\boxed{\textit{if } A, \textit{ then } B}$ (the double negation is eliminated). All participants responded correctly that they are pretty certain that the conclusion is true.

The probabilistic truth table tasks All participants of Experiment 2 ($n_3+n_4 = 39$) solved four probabilistic truth table tasks. This task serves to investigate the representation of uncertain conditionals. The difference to the other tasks is that the probabilistic truth table task investigates reasoning about complete probabilistic knowledge. We adapted the tasks from Evans et al. (2003). The participants were instructed to imagine a pack of 120 cards. They were informed that the cards are either red or blue and have have either a circle or a square printed on them. They were asked to assume that in total there are:

40 red and circle,
40 red and square,
20 blue and circle, and
20 blue and square cards.

The participants were instructed to imagine that the pack is shuffled and that a card is drawn randomly. This makes the random process explicit. Then, they rated the following four assertions about the randomly drawn card:

(1) **If** the card is blue, **then** there is a circle on it.

(2) **If** there is a square on the card, **then** it is red.

(3) **If** the card is **not** red, **then** there is **not** a square on it.

(4) **If** there is a **not** a circle on the card, **then** it is **not** blue.

For the ratings we provided scales with the labels “absolutely certain **not** the case” (German original: stimmt absolut sicher **nicht**) and “absolutely certain the case” (stimmt absolut sicher).

The most important qualitative predictions are as follows. (1) and (3) are equivalent, as are (2) and (4). If the participants interpret the “if—then—” as a material conditional, all four assertions should obtain the same rating. If the participants interpret the “if—then—” as a conditional event or as a conjunction, then (1) and (3) should obtain a lower rating than (2) and (4), respectively. For differentiating between the conditional event and the conjunction interpretation, the ratings must be compared with the normative values. The quantitative predictions are summarized in Table 7. The predictions according to the conditional event interpretation are higher than the predictions according to the conjunction interpretation.

Conditional	Predictions			Responses
	$P(\cdot \supset \cdot)$	$P(\cdot \wedge \cdot)$	$P(\cdot\|\cdot)$	Mean (SD)
If blue, then circle	.83	.17	.50	.46 (.18)
If square, then red	.83	.33	.67	.71 (.15)
If not red, then not square	.83	.17	.50	.44 (.22)
If not circle, then not blue	.83	.33	.67	.66 (.23)

Table 7: Predicted probabilities and mean observed values in the four probabilistic truth table tasks of Experiment 2 ($n = 39$). The predicted probabilities are the normative probabilities according to the material conditional ($\cdot \supset \cdot$), conjunction ($\cdot \wedge \cdot$), and according to the conditional event ($\cdot|\cdot$) interpretation of uncertain indicative conditionals.

We counted the ratings as equal if they remained in a 15% interval. Only two of the thirty-nine participants responded with equal ratings for all four assertions. This is clear evidence against the material conditional interpretation of indicative conditionals. Thirty participants rate the assertions (2) and (4) uniformly higher than both assertions (1) and (3). This pattern is predicted by both, the conjunction interpretation and the conditional event interpretation. Twenty-five participants responded by giving *both* (i) equal ratings between the assertions (1) and (3), and (ii) equal ratings between the assertions (2) and (4) (within a 15% interval). This shows that most participants really understood that the respective assertions are equivalent.

We divided the distance between the left pole of the rating scale ("absolutely certain **not** the case") and the participants' markings by the total length (63mm) of the response scale. This procedure scales the response values from zero to one. Table 7 presents the mean response values and the normative probabilities of the material conditional, conjunction, and conditional event interpretation of the "if—, then—". Boxplots of the data are given in Figure 1. The mean and the median of the responses are close to the predictions of the conditional event interpretation. Therefore, conditional probability is the best predictor of the empirical values. We observed fewer responses that are consistent with the conjunction interpretation than reported in the literature (see, e.g., Evans et al., 2003; Oberauer & Wilhelm, 2003). A reason could be that we maximized the distances between the three predictions, and the coherent value of the conjunction interpretation is quite far away form the coherent value of the conditional event interpretation.

Moreover, we measured the distances between each response and each of the three normative predictions. The least distances provide another way of comparing the quality of the three predictions. A prediction "wins" if it has the least distance to the response value compared with the other predictions. Table 8 presents the frequencies of the "winning" predictors. The conditional event interpretation received the highest frequencies. Again, the conditional event is the best predictor of the empirical values.

6 Discussion

The paradoxes of the material conditional are often used as one of the principal arguments for why one should not interpret indicative conditionals, "if A, then B", as material conditionals, $A \supset B$. However, to our knowledge the paradoxes have not been investigated empirically yet. This contribution reports two first experiments on the paradoxes of the material conditional. One main result is that the great majority of participants do not endorse the paradoxes. We proposed *coherence based probability logic* as a normative framework for the psychology of reasoning. Within this framework we formulated

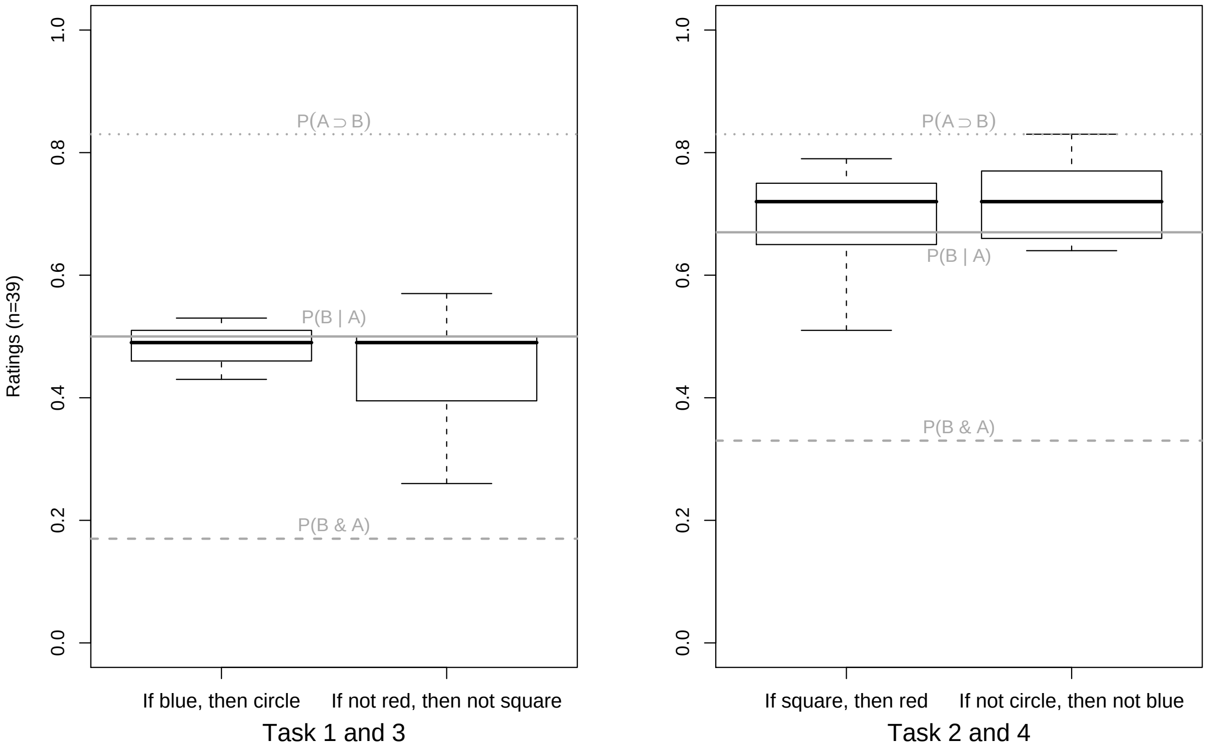

Figure 1: Boxplots of the probabilistic truth table tasks. The boxes contain 50% of the responses and the median (thick line). The whiskers indicate 1.5 $\times$ the interquartile range. The three normative predictions are in gray. Conditional probability is the best predictor.

Conditional	Winning predictor			Ties
	$P(\cdot \supset \cdot)$	$P(\cdot \wedge \cdot)$	$P(\cdot\|\cdot)$	
If blue, then circle	2	7	30	0
If square, then red	9	4	22	4
If not red, then not square	3	9	27	0
If not circle, then not blue	12	8	18	1

Table 8: Frequencies of least distances to the normative predictors ("winning predictor"; $n_3 + n_4 = 39$). "Ties" denotes the number of participates that responded by values that lie exactly between two predictors. Ties were observed only between $P(\cdot \supset \cdot)$ and $P(\cdot|\cdot)$.

three psychologically prominent interpretations of the "if—, then—", and explained why most people do not endorse the paradoxes: most people interpret the "if—, then—" as a conditional event and they seem to understand that the paradoxes are *probabilistically non-informative*. Pragmatic reasons may play important rôles in human reasoning in general, but for the paradoxes probabilistic non-informativeness fully explains why most people do not endorse these argument forms. We provided further empirical support for the conditional event interpretation by the results of the probabilistic truth table task, which endorses Jonathan Evans' support of the conditional probability interpretation of uncertain conditionals (Evans & Over, 2004; Evans, 2007). We explained why all this evidence rejects central predictions of the mental model theory.

We started from a strong assumption, namely that there is just *one* logical system that fits all subjects, all tasks, all situations and, moreover, all their combinatorial interactions. This assumption is too strong. There are, for example, two subjects who give exactly equal ratings to all four truth table tasks. This is what the material conditional of classical logic predicts. *Most* subjects, though, follow the predictions of probability logic. It is plausible to assume that there are not only systematic individual differences but that even one and the same subject has access to several different "logics" of reasoning. Human thinking can be excellent in switching strategies, in simplifying constraints, or in re-representing a problem structure. We have tried to identify and discuss one dominating system, probability logic.

Finding moderator variables that tell us who in which situation is using which kind of "logic" to make which inference is difficult. The large number of variables, the small effect sizes, and the realization of appropriate experimental conditions for motivated "real" thinking (and not just giving the next best guess answers) make empirical research difficult and require considerable efforts.

Is our position not a kind of logicism? Do we not say that to come up with a theory of human reasoning, we need to identify the formal system that best fits human performance?

Logic and mathematical computer science provide the best theories on reasoning, not about human reasoning, of course. But they provide languages, properties, and systems to describe tasks, knowledge and belief, and inferences that the psychology of reasoning cannot ignore.

Ignoring modern logics leads to the use of a common sense amateur logic. There is no way out, either (i) explicitly exploiting the modern pluralistic approaches to logic, or (ii) relying on classical logic, or (iii) implicitly using self-baked logic. For about one hundred years parts of classical logic served as the favorite reference system to select experimental tasks and to build models. Typical examples are syllogisms, the Wason's card selection task, or the MODUS PONENS, MODUS TOLLENS, AFFIRMING THE CONSEQUENT, and the DENYING THE ANTECEDENT tasks. The research questions had been centered around the logic of syllogisms (a highly specific subset of predicate logic going back to Aristotle and the middle ages), a special form of the conditional (the so-called material conditional), and just four forms of conditional syllogisms with two premises. Only recently systems that differ from classical logic were considered for psychological modeling (and for a re-analysis of the classical tasks). We are not saying that cognitive representations and processes are unimportant. Quite the contrary, cognitive representations and processes are of course central to psychological theory building. We have severe doubts, however, that it is a good research strategy to attempt to describe representations and processes on the background of an old-fashioned amateur logic. Moreover—and again—logic and mathematical computer science offer a rich repertoire of knowledge representation languages and systems. Several well-known representational systems in psychology were imported from computer science. Semantic networks, neural networks, production systems, or more recently, Bayesian networks are typical examples.

References

Adams, E. W. (1975). *The logic of conditionals.* Dordrecht: Reidel.

Antoniou, G. (1997). *Nonmonotonic reasoning.* Cambridge: MIT Press.

Baader, F., Calvanese, D., McGuinness, D. L., Nardi, D., & Patel-Schneider, P. F. (Eds.). (2007). *The description logic handbook: Theory, implementation, and applications.* Cambridge: Cambridge University Press.

Barrouillet, P., & Lecas, J.-F. (1999). Mental models in conditional reasoning and working memory. *Thinking and Reasoning*, *5*(4), 289-302.

Biazzo, V., Gilio, A., Lukasiewicz, T., & Sanfilippo, G. (2005). Probabilistic logic under coherence: Complexity and algorithms. *Annals of Mathematics and Artificial Intelligence*, *45*(1-2), 35-81.

Coletti, G., & Scozzafava, R. (2002). *Probabilistic logic in a coherent setting.* Dordrecht: Kluwer.

Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement.* Hove, UK: Psychology Press.

Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, *11*, 295-306.

Evans, J. St. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(2), 321-355.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning.* Hove: Erlbaum.

Evans, J. St. B. T., & Over, D. E. (2004). *If.* Oxford: Oxford University Press.

Gilio, A. (2002). Probabilistic reasoning under coherence in System P. *Annals of Mathematics and Artificial Intelligence*, *34*, 5-34.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*(4), 646-678.

Lewis, C. I. (1918). *A survey of symbolic logic.* Berkeley: University of California Press.

Liu, I.-M. (2003). Conditional reasoning and conditionalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(4), 694-709.

Liu, I.-M., Lo, K.-C., & Wu, J.-T. (1996). A probabilistic interpretation of 'If—Then'. *The Quarterly Journal of Experimental Psychology*, *49*(A), 828-844.

Lukasiewicz, T. (2005). Weak nonmonotonic probabilistic logics. *Artificial Intelligence*, *168*, 119-161.

Lukasiewicz, T. (2008). Expressive probabilistic description logics. *Artificial Intelligence*, *172*(6-7), 852 - 883.

Oaksford, M., & Chater, N. (Eds.). (1998). *Rational models of cognition.* Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2007a). *Bayesian rationality: The probabilistic approach to human reasoning.* Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2007b). *Bayesian rationality: The probabilistic approach to human reasoning.* Oxford: Oxford University Press.

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 883-899.

Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional probabilities, mental models and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 680-693.

Over, D. E., & Evans, J. St. B. T. (2003). The probability of conditionals: The psychological evidence. *Mind & Language*, *18*, 340-358.

Pfeifer, N., & Kleiter, G. D. (2003). Nonmonotonicity and human probabilistic reasoning. In *Proceedings of the 6th workshop on uncertainty processing* (p. 221-234). Hejnice: September 24–27$^{\text{th}}$, 2003.

Pfeifer, N., & Kleiter, G. D. (2005a). Coherence and nonmonotonicity in human reasoning. *Synthese*, *146*(1-2), 93-109.

Pfeifer, N., & Kleiter, G. D. (2005b). Towards a mental probability logic. *Psychologica Belgica*, *45*(1), 71-99.

Pfeifer, N., & Kleiter, G. D. (2006). Inference in conditional probability logic. *Kybernetika*, *42*, 391-404.

Pfeifer, N., & Kleiter, G. D. (2007, 16 - 19 July). Human reasoning with imprecise probabilities: Modus ponens and Denying the antecedent. In G. De Cooman, J. Vejnarová, & M. Zaffalon (Eds.), *5th International Symposium on Imprecise Probability: Theories and Applications* (p. 347-356). Prague, Czech Republic.

Pfeifer, N., & Kleiter, G. D. (in pressa). The conditional in mental probability logic. In M. Oaksford (Ed.), *The psychology of conditionals.* Oxford: Oxford University Press.

Pfeifer, N., & Kleiter, G. D. (in pressb). Framing human inference by coherence based probability logic. *Journal of Applied Logics*.

Van Harmelen, F., Lifschitz, V., & Porter, B. (Eds.). (2008). *Handbook of knowledge representation.* Amsterdam: Elsevier.

Wiles, A. (1995). Modular elliptic curves and Fermat's Last Theorem. *Annals of Mathematics*, *142*, 443-551.

Acknowledgments

This work is supported by the LogICCC project "*The Logic of Causal and Probabilistic Reasoning in Uncertain Environments*" (European Science Foundation/Austrian Research Fonds) and by the FWF project "*Mental probability logic*" (Austrian Research Fonds). The authors thank Leonhard Kratzer for collecting the data, and David Over and Andy Fugard for helpful comments.

The Conditional in Mental Probability Logic

Pfeifer, N. & Kleiter, G. D.
Department of Psychology, University of Salzburg (Austria)
niki.pfeifer@sbg.ac.at, gernot.kleiter@sbg.ac.at

(December 29, 2007)

1 Introduction

Since Störring's [63] pioneering experiments on syllogistic reasoning at the beginning of last century, experimental psychology has investigated deductive reasoning in the framework of classical logic. The most prominent examples are the theories of mental models [27] and mental rules [8, 59]. A fragment of the model theory of classical logic is central to mental models. Likewise, a fragment of the proof theory of classical logic is central to mental rules. In this tradition, classical logic is considered as the "surest guide" towards a competence model for the psychology of reasoning [38]. Not only did classical logic guide the psychological theories, but it also determined the experimental methodology, and the evaluation of human performance.

In the last decade the situation has changed. At present, approaches that extend, or go beyond, classical logic introduce new frameworks in the field. Examples are nonmonotonic reasoning, possibility theory [3], logic programming [61, 62], probabilistic approaches [41, 11, 43, 42, 37, 36, 19, 46, 47, 44] ... [links to other chapters in the book].

The present chapter describes a probabilistic framework of human reasoning. It is based on probability logic. While there are several approaches to probability logic, we adopt the *coherence* based approach [13, 23]. We assume that rules similar to the principles of probability logic are basic rules of the human inference engine. We therefore call our approach "mental probability logic" [51]. Conditionals are of special importance in the approach. Their interpretation is different from the interpretation in other approaches. We conceive conditionals as non-truth functional, as uncertain, and as nonmonotonic. They allow for exceptions. Below, we call such conditionals "nonmonotonic conditionals". We note that causal, counterfactual, deontic, or pragmatic conditionals [5] are not in the scope of this chapter, because their logical forms require formalisms that go beyond the scope of the present framework. Causal conditionals require logical operators for intervention, counterfactuals and deontic conditionals require possible worlds semantics, and pragmatic conditionals require a theory of the context of their uttering.

Many investigations on cognitive processes report errors, fallacies, or biases. Well known are perceptual illusions, biases in judgment under uncertainty, or errors in deductive reasoning. While these phenomena may be startling and stimulating in the scientific process, they do not lead to theories that explain human performance in a *systematic* way. Collecting slips of the tongue does not lead to a theory of speaking. Such phenomena should be integrated in a systematic theory and not studied in isolation.

Psycholinguists distinguish performance and competence [12]. Competence describes what functions a cognitive system can compute [39, 40]. Human reasoning can solve complex problems and perform sophisticated inferences. When developing a theory of reasoning one should have an explanation of these processes in mind. Such an explanation requires a competence theory on the on the computational level. In the long run, we want to develop a psychological theory that accounts for both competence and performance. At the competence level, a systematic formal theory is required. The formal theory provides the rationality standards and provides tools for computational descriptions of the human reasoning competency. At the performance level a specification of the cognitive representations and processes is required. The explanation of typical reasoning, good and bad inferences requires a theory of how representations are formed and manipulated.

On the competence level, classical logic provided a rich systematic framework. Nonmonotonic reasoning systems, like SYSTEM P [32, 1, 23], provide a more promising framework. For several reasons classical logic alone is a Procrustean framework [40, 42, 51, 52]. The two most important reasons are the monotonicity property (i) and the IF–THEN relation (ii).

(i) Monotonicity is a meta-property of classical logic. It states that adding premises to a valid argument can only increase the set of conclusions. Monotonicity does not allow to retract conclusions in the light of new evidence. In everyday life, however, we often retract conclusions when we face new evidence. Moreover, experiments on the suppression of conditional inferences show that human subjects withdraw conclusions when new evidence is presented [9, 10, 6, 7, 16, 57]. Thus, the monotonicity principle is psychologically implausible. We discuss a coherence based semantic for nonmonotonic reasoning and empirical results below.

(ii) The conditional in classical logic is the *material conditional*. Table 1 lists its truth conditions. The material conditional, $A \supset B$, is true if, and only if, it is not the case that the *antecedent*, A, is true and the *consequent*, B, is false.

While the material conditional is extremely useful in formal fields like mathematics (derivations, proofs), it has severe drawbacks in the formalization of common sense conditionals. In common sense reasoning, conditionals are inherently uncertain, as they hold only "probably", "normally", or "usu-

State of world A	B	*Material conditional* $A \supset B$	*Betting interpretation* $B\|A$
t	t	t	win
t	f	f	lose
f	t	t	money back
f	f	t	money back

Table 1: Truth table of the material conditional, and the betting interpretation of the conditional event, $B|A$. "A" and "B" denote propositions. "t" and "f" denote "true" and "false", respectively.

ally". A few exceptions do not invalidate the conditional. Nonmonotonic conditionals express uncertain relations between the IF and the THEN part of a conditional assertion. The nonmonotonic conditional is interpreted as a "high" conditional probability assertion,

$$\boxed{\textit{If } A \textit{ normally } B} \text{ if, and only if,}$$
$$\boxed{\textit{the probability of } B \text{ given } A,\ P(B|A),\ \textit{is "high"}} \quad .$$

Here, the probability function, $P(\cdot)$, is a one-place function and the conditional event, $B|A$, is its argument. The conditional event, $B|A$, is distinct from the material conditional of logic, $A \supset B$. In the following paragraphs we argue why the core of the IF–THEN corresponds to the conditional event and why it does not correspond to the material conditional.

The material conditional leads to counterintuitive consequences, known as the paradoxes of the material conditional. Below, we discuss an empirical study on one of these paradoxes (PREMISE STRENGTHENING[1]). We do not want to import the paradoxes of the material conditional into the mental probability logic. This is one reason why we interpret the non-probabilistic conditional as a conditional event, $B|A$ [1, 52]. The paradoxes arise because of the truth-functionality of the material conditional, which will be discussed in the next paragraph.

The truth value of the material conditional is determined for all four possible combinations of truth values of the antecedent and the consequent (see Table 1). Therefore, the material conditional is truth functional. In the

[1] PREMISE STRENGTHENING is an argument with one premise (first box) and one conclusion (second box): *from* $\boxed{A \supset B}$ *infer* $\boxed{(A \wedge C) \supset B}$. Can the conditional in the conclusion be false if the premise is true? If $A \wedge C$ is false, then the conditional is true (because false antecedents make the material conditional true). If $A \wedge C$ is true, the conditional is true (because of the premise). Thus, it cannot be the case that the premise is true and the conclusion is false at the same time. Therefore, PREMISE STRENGTHENING is logically valid.

long history of logic,[2] the truth functionality of the material conditional was criticized several times, especially for those cases in which the antecedent is false. It is counter the intuition to call a conditional true if its antecedent is false. PREMISE STRENGTHENING, for example, is logically valid because (per definition) the material conditional is true if its antecedent is false. Ramsey [58] and de Finetti [14, 15] pointed out that the truth value of the conditional event, $B|A$, is *indeterminate* if the conditioning event, A, is false. In a betting interpretation this corresponds to the case in which a bet is annulled if the conditioning event does not happen. If you bet, for example, that

team X wins

on the condition that

team X plays against team Y ,

then the stakes are payed back in the case that the game is cancelled (*team X does not play against team Y*; see Table 1). If the conditioning event does not happen, the conditional event is not true. The conditional event is indeterminate if the conditioning event is false. Thus, the vertical stroke | in the conditional event is not truth functional. Therefore, the paradoxes of the material conditional do not arise.

The conditional event is a *genuine* conditional. It cannot be constructed by Boolean operators like negation ($\neg$, "not"), disjunction ($\vee$, "or"), or conjunction ($\wedge$, "and"). The material conditional, however, is not a genuine conditional. The definition of the material conditional depends upon which Boolean operator is considered to be basic. $A \supset B$ can be defined, for example, by negation and disjunction, negated conjunction, or by intuitively indigestible definientia.[3] None of these definientia are conditionals, but they are logically equivalent to the material conditional. A genuine conditional (like the nonmonotonic conditional) cannot be expressed by something that goes completely beyond IF–THEN formulations. Therefore, we prefer genuine conditionals to non-genuine conditionals.

We see that the conditional probability interpretation of nonmonotonic conditionals has at least three theoretical advantages compared with the material conditional interpretation: (i) *probability* accounts for uncertainty and nonmonotonicity, (ii) *conditional events* are genuine conditionals and (iii) conditional events are free of the paradoxes of the material conditional. What is the empirical status of the conditional probability interpretation? The next section gives a brief overview on recent probabilistic approaches to human conditional reasoning.

[2]The roots of the material conditional go back to Philon of Megara. He lived around the 4th and 3rd century BC [31].

[3]For example, $A \supset B$ is definable by negation and disjunction $\neg A \vee B$, by the negated conjunction $\neg(A \wedge \neg B)$, and by the intuitively indigestible definiens $((A \downarrow A) \downarrow B) \downarrow ((A \downarrow A) \downarrow B)$ as well, where "$A \downarrow B$" is read as "neither A, nor B" ($\neg(A \vee B)$).

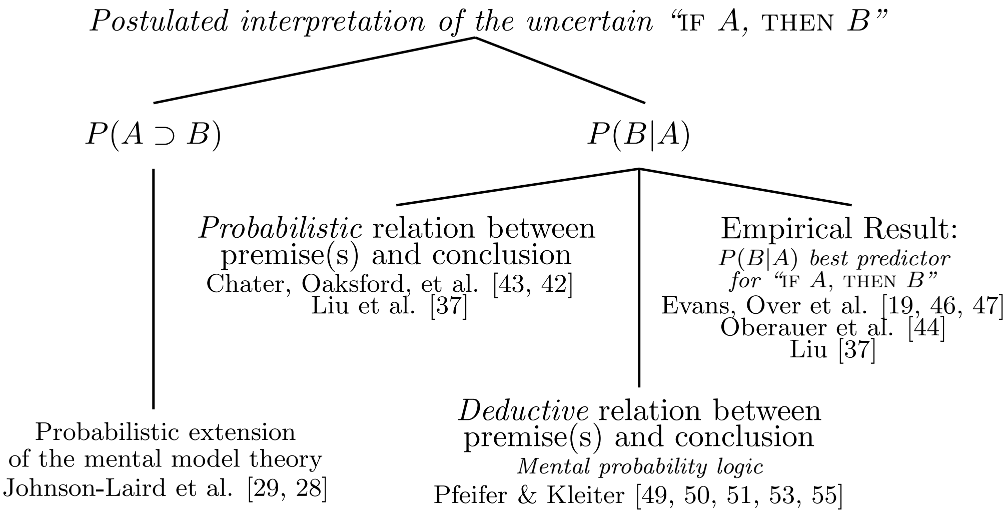

Figure 1: Probabilistic approaches to human conditional reasoning and selected exemplar studies. $P(A \supset B)$ denotes the probability of the material conditional. $P(B|A)$ denotes the conditional probability.

2 Probabilistic approaches to human conditional reasoning

There is a long tradition of probabilistic approaches to human judgment and decision making. The judgment/decision tasks were associated with inductive reasoning. Therefore the heavy use of probability theory. In the traditional psychology of reasoning, however, the tasks were associated with deductive reasoning. Therefore the heavy use of classical logic. Recently, both traditions began to merge [30]. In the early nineties of the last century, Chater and Oaksford introduced probabilistic models to the field of deductive reasoning [40, 41, 11, 43, 42]. The probabilistic approach to deductive reasoning claims that even "purely logical" tasks are solved as if they were tasks belonging to probability theory. The Wason Selection Task, for example, is solved as if the human subjects would maximize information gain [41]. Syllogisms are solved as if the subjects would process Bayesian probabilities [11]. And argument forms like the MODUS PONENS or the MODUS TOLLENS are solved as if the subjects were experts in probability logic [55, 56]. During the last five years, the interest in probabilistic approaches to human deductive reasoning increased rapidly. Recent probabilistic approaches to human deductive reasoning can roughly be classified by (i) the postulated interpretation of the IF–THEN and (ii) the postulated relation between the premise(s) and the conclusion (see Figure 1).

The truth conditions of the IF–THEN in the mental model theory coincide with the material conditional (left branch of Figure 1). If the subjects "fully flesh out" all truth table cases of the IF–THEN, then P(IF–THEN) is equal to the probability of the material conditional (see [28, 21] for a discussion).

On the right hand side of Figure 1 is the conditional probability interpretation (the IF–THEN is interpreted as a conditional event). We note that the numerical probabilities of the material conditional, $P(A \supset B)$, and of the corresponding conditional probability, $P(B|A)$, can differ substantially. Dorn [17] gives a compelling example. Consider a fair die with six sides. Let A be *the next throw of the die will come up a 5*, and let B be *the next throw of the die will come up an even number*. By the way, IF A, THEN B is here intuitively implausible. A denotes one out of the six possible outcomes, thus $P(A) = 1/6$. B denotes three out of the six possible outcomes, thus $P(B) = 1/2$. Summing up the probabilities of those truth table cases that make $A \supset B$ true (see Table 1), gives $P(A \supset B) = 5/6$. $P(B|A)$ is determined only if a 5 comes up (if a 5 does not comes up, A is false). If A is true, then B is false (because 5 is not an even number), hence $P(B|A) = 0$. $P(B|A) = 0$ reflects the fact that here IF A, THEN B is intuitively implausible.

As the values of $P(A \supset B)$ and of $P(B|A)$ can differ substantially, you might ask, which of both interpretations predicts the human understanding of IF–THEN better? Studies on human understanding of IF–THEN do not endorse the probability of the material conditional interpretation. Rather, conditional probability seems to be the best predictor for human understanding of IF–THEN [37, 19, 46, 47, 44]. We take this result as an important building block for a competence theory of human reasoning. In the following sections, we discuss human inference from conditional premises.

3 Coherence based probability logic

One of the best known principles in probability logic is Adams' concept of p-validity [1, 2]. An argument is p-valid, if and only if, the uncertainty of the conclusion of the argument cannot exceed the sum of the uncertainties of its premises. The uncertainty $u(A)$ is defined by the 1-complement of the corresponding probability, $u(A) = 1 - P(A)$. If an argument is p-valid, then the corresponding non-probabilistic argument is logically valid. Logical validity, however, does not guarantee p-validity.

In terms of interval probability, the lower probability of a p-valid conclusion is not sensitive (i) to the specific logical form of the premises and (ii) to the order of the probabilities of the premises. The two properties hold for unconditional events only and reflect the fact that in this case the events are truth functional. Only the lower bounds of the conclusions of those arguments that contain conditional events can be sensitive to the structure of the premises and to the specific pattern of the probability assessment.

If human subjects interpret the IF–THEN as a material conditional, then their probability responses in p-valid arguments should be insensitive (i) to the logical form of the premises and (ii) to permutations of the probabilities of the premises. There is, however, strong evidence that *human subjects are sensitive to structure and assignment.*

We think that the investigation of lower and upper probabilities is important for the psychology of reasoning. We investigate structure, assignment, and inference in probabilistic argument forms. If the probability of the conclusion is constrained by the probabilities of the premise(s), the inference is called "probabilistically informative". If the assignment of the unit interval, $[0, 1]$, to the conclusion is coherent under *all* assessments of the premise(s), the inference is called "probabilistically not informative". In this case the premises do not constrain the probability of the conclusion. As a trivial example, assume you know that $P(A) = .7$. Based on this premise, you can only infer that $P(B) \in [0, 1]$. This is probabilistically not informative.

While logical validity is a necessary condition for p-validity, logical validity is not a necessary condition for probabilistic informativeness. The non-probabilistic forms of the DENYING THE ANTECEDENT[4] and AFFIRMING THE CONSEQUENT[5] are not logically valid, but the probabilistic versions are probabilistically informative (but not p-valid). Moreover, PREMISE STRENGTHENING, HYPOTHETICAL SYLLOGISM[6], and CONTRAPOSITION[7] are logically valid, but neither probabilistically informative nor p-valid.

If the premises of a probabilistically informative argument are certain (probabilities equal to 1), and if the argument is logically valid, then its conclusion is certain. If the premises of a probabilistically informative argument are certain, and if the argument is not logically valid, then the probability of its conclusion may be anywhere in the unit interval. If all premises are given for sure, then the logically invalid arguments make also probabilistically no sense. Classical logic is thus a "limiting case" for probabilistically informative arguments.

This special role of classical logic is an important reason why we do not want to exclude classical logic from our approach. In everyday life, however, premises are usually not given for sure. In these cases classical logic does not provide an appropriate theoretical frame. Probabilistic versions of argument forms and the relationships between logical validity, probabilistic informativeness, and p-validity are investigated in [52, 56].

In psychology, the most often investigated argument forms containing conditionals are the conditional syllogisms MODUS PONENS and MODUS TOLLENS, and the related logical fallacies DENYING THE ANTECEDENT and AFFIRMING THE CONSEQUENT. Each of the four argument forms consists of the conditional premise IF A, THEN B, one categorical premise, and a con-

[4] *from* $\boxed{\text{IF } A \text{ THEN } B \textit{ and not-}A}$ *infer* $\boxed{\textit{not-}B}$

[5] *from* $\boxed{\text{IF } A \text{ THEN } B \textit{ and } B}$ *infer* $\boxed{A}$

[6] *from* $\boxed{\text{IF } A \text{ THEN } B \textit{ and } \text{IF } B \text{ THEN } C}$ *infer* $\boxed{\text{IF } A \text{ THEN } C}$

[7] *from* $\boxed{\text{IF } A \text{ THEN } B}$ *infer* $\boxed{\text{IF } \textit{not-}B \text{ THEN } \textit{not-}A}$

clusion. In the original probabilistic approach of Oaksford, Chater & Larkin [43] the probability of the conclusion of a conditional syllogism is equal to the conditional probability of the conclusion given the categorical premise. As an example consider the MODUS PONENS,

$$\textit{from}\ \overbrace{\underbrace{\text{IF } A, \text{ THEN } B}_{\textit{conditional}}\ \textit{and}\ \underbrace{A}_{\textit{categorical}}}^{\textit{Premises}}\ \textit{infer}\ \overbrace{B}^{\textit{Conclusion}}\ .$$

$P(B|A)$ predicts the endorsement of the MODUS PONENS. The conditional premise, IF A, THEN B, is ignored in the model. The original model was modified [36, 42] such that the conclusion is conditionalized on the categorical and on the conditional premise. In this approach, the inference-relation between the premise(s) and the conclusion is uncertain (see Figure 1).

Our approach follows a different intuition (see Figure 1). We assume a coherent probability assessment of the premise(s) and the inference problem consists in deriving *deductively* the (interval-)probability of the conclusion. Elementary probability theory provides rules how to deduce the probability of a target event (the conclusion) from the probabilities of a number of other events (the premise(s)). In general, we consider as *premises* a triple consisting of (i) a given set of arbitrary *conditional events*, $A_1|B_1, \ldots, A_n|B_n$, (ii) the associated probability assessment $p_1, \ldots, p_n$, and (iii) a (possibly empty) set of logical relations between the events.[8] The *conclusion* is a further conditional event $A_{n+1}|B_{n+1}$. The inference problem is solved when $P(A_{n+1}|B_{n+1})$ is determined. Bayes' Theorem, for example, finds the probability of a conditional event, $A|B$, when the probabilities of three other events, A, $B|A$, and $B|\neg A$, are given. Bayes' Theorem may then be written as an inference rule

$$\begin{array}{ll} \textit{from} & \boxed{P(A) = x \textit{ and } P(B|A) = y \textit{ and } P(B|\neg A) = z} \\ \textit{infer} & \boxed{P(A|B) = xy/(xy + (1-x)z)}\ , \end{array}$$

where the first box contains the premises (the assessment is assumed to be coherent) and the second box contains the conclusion.

To evaluate the rationality of human inferences, we investigate to what extent humans infer *coherent* probability assessments from a given set of premises. A probability assessment is *coherent* if it does not admit one or more bets with sure loss (often called a "Dutch book"). Compared with the criterion of maximizing expected utility (traditionally used in the judgment and decision making literature), coherence is much weaker. Coherence is one of the key concepts in the theory of subjective probability. It was introduced

[8] *Unconditional events* are treated as special cases of conditional events. An unconditional event, A, is defined as the conditional event A *given logical truth*, $A|verum$. The according probabilities are identical, $P(A) =_{def.} P(A|verum)$. "$P(A)$" is a shortcut for "$P(A|verum)$".

by de Finetti [14, 15]. More recent work includes [64, 33, 13, 23]. Coherence provides an adequate normative foundation for the mental probability logic and has many psychologically plausible advantages compared with classical concepts of probability:

- Coherence is in the tradition of subjective probability theory in which probabilities are conceived as *degrees of belief*. Degrees of belief are coherent descriptions of partial knowledge states. For the mental probability logic framework, the interpretation of probability as degrees of belief is naturally more appropriate than "relative frequency" interpretations of probability (for example, Reichenbach, or von Mises). Relative frequency interpretations of probability are about "objective entities" in the outside world. Mental probability logic, however, investigates epistemic states of uncertainty.

- The framework of coherence does not require to start from a *complete Boolean algebra*. Complete algebras are psychologically unrealistic since they can neither be unfolded in the working memory nor be stored in the long term memory. Humans try to keep the memory load as small as possible and try to process only relevant information. Only the information contained in the premises is relevant for drawing inferences. Additional probabilities constitute additional premises.

- *Conditional probability, $P(B|A)$,* is a *primitive* notion. The probability value is assigned *directly* to the conditional event, $B|A$, as a *whole* [13]. The conditioning event A must not be a logical contradiction, but $P(A)$ can be equal to zero. The method of assigning the probability values directly to the conditional event, $B|A$, contrasts with the classical approach to probability, where conditional probability is defined by the fraction of the "joint", $P(A \wedge B)$, and the "marginal", $P(A)$, probabilities, where $P(A) \neq 0$. It is psychologically plausible that subjects usually assign the uncertainty directly to the conditional (and not by building fractions).

- Because of lack of knowledge (time, effort), it may be impossible for a person to assign precise probabilities to an event. *Imprecise* probability assessments my be expressed by interval-valued probabilities or by second order probability distributions [54].

These advantages explain why we take coherence and not the classical concept of probability as the normative basis for the mental probability logic. The subsequent sections summarize selected empirical work, and we discuss to which extent coherence based probability logic describes actual human inferences. Studies on the conditional syllogisms are reported in [55, 56]. Studies on the nonmonotonic SYSTEM P rules, and argument forms that

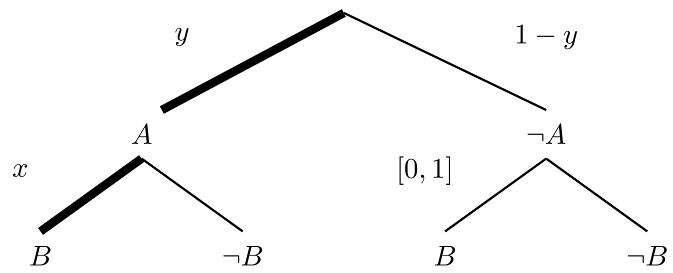

Figure 2: Derivation of the probabilistic MODUS PONENS. The probabilities of the premises are given, $P(B|A) = x$ and $P(A) = y$ (the bold branches on the left). The lower and the upper probability bounds of the conclusion, $P(B)$, are derived by the theorem of total probability, $P(B) = P(A)P(B|A) + P(\neg A)P(B|\neg A)$. $P(B|\neg A)$ is unknown and can take any value in $[0, 1]$. $P(B) = xy$ if $P(B|\neg A) = 0$, and $P(B) = xy + 1 - y$ if $P(B|\neg A) = 1$. Therefore, $xy \leq P(B) \leq xy + 1 - y$.

are monotonic counterparts of the SYSTEM P rules (HYPOTHETICAL SYLLOGISM, PREMISE STRENGTHENING, and CONTRAPOSITION) are reported in [49, 50, 53].

3.1 Example 1: The nonmonotonic conditional in the modus ponens

3.1.1 Formal background

The logical form of the MODUS PONENS is

$$\textit{from } \boxed{A \supset B \textit{ and } A} \textit{ infer } \boxed{B}\ ,$$

where A and B denote propositions and $\supset$ denotes the material conditional. The MODUS PONENS is logically valid. The conditional probability version of the MODUS PONENS is

$$\textit{from } \boxed{P(B|A) = x \textit{ and } P(A) = y} \textit{ infer } \boxed{P(B) \in [z', z'']}\ ,$$

where the probability value of the conditional premise is equal to x and the probability value of the categorical premise is equal to y. The probability of the conclusion (B) is in the probability interval from *at least* z' to *at most* z'', $[z', z'']$. The derivation of the coherent probability interval of the conclusion of the MODUS PONENS is explained in Figure 2. The lower bound z' is equal to the product xy and the upper bound z'' is equal to $xy + 1 - y$. Thus, the probability of the conclusion is constrained by the probabilities of the premises. This argument form is also p-valid.

The uncertainty of the premises may be expressed by interval-valued probabilities. A person may specify, for example, that an event A has at least probability x. The MODUS PONENS with interval probabilities in the premises has the form

$$\begin{array}{ll} \textit{from} & \boxed{P(B|A) \in [x', x''] \ \textit{and} \ P(A) \in [y', y'']} \\ \textit{infer} & \boxed{P(B) \in [x'y', 1 - y' + x''y']} \ , \end{array}$$

where x' and x'' are the lower and upper bounds of the conditional premise, and y' and y'' are the lower and upper bounds of the categorical premise, respectively. If a person knows a lot about the propositions in the premises, then she will assess tight intervals. If her knowledge is vague and ambiguous, then she will assess wide intervals.

Imprecise probabilities are sometimes criticized by the following argument. It is paradox to say that, if a person is not able to assess one precise point probability, she may overcome the difficulty by assessing now two precise probabilities. We only partially agree with this argument. In everyday life intervals are very often used to communicate imprecision. We would prefer to represent degrees of belief by distributions which are "smeared" across the whole zero-one range. Statistics uses second order probability density functions to describe knowledge about uncertain quantities. This complicates the formal models considerably though. In the present context it seems reasonable to consider interval probabilities as approximations to confidence intervals. It is possible, however, to replace "probability logic" by a "statistical logic" which investigates logical argument forms with probability distributions. We described first steps in [54]. An advantage of such an approach is the possibility to update the distributions in the light of observational data like frequencies or averages. Bayesian statistics offers a rich theoretical background.

3.1.2 Empirical investigation of the modus ponens

In our experiments, we try to construct cover-stories that have a neutral content, that is as independent as possible from the background knowledge of the subjects. Moreover, we take care that only the information explicitly contained in the argument enters the task. The MODUS PONENS, for example, involves only two premises. Accordingly, the probabilistic version of the MODUS PONENS contains only two probabilities, $P(B|A)$ and $P(A)$. We translated the probabilistic MODUS PONENS into several cover-stories, of the following kind [55]:

> Claudia works at the blood donation services. She investigates to which blood group the donated blood belongs and whether the donated blood is Rhesus-positive.
>
> > Claudia is 60% certain: <u>If</u> the donated blood belongs to the blood group A, <u>then</u> the donated blood is Rhesus-positive.
> > Claudia is 75% certain: A recent donated blood belongs to the blood group A.
>
> How certain should Claudia be that this recent donated blood is Rhesus-positive?

The cover-stories contained the probabilities of the premises. The task of the participants was to infer from the premises the probability(-interval) of the conclusion. In all experiments we paid special attention to encourage the participants to engage in reasoning and to avoid quick guessing. The participants were students of the Salzburg University. They were tested individually in a quiet room in the department. They were asked to take enough time.

Introductory examples explained that the solution can be a point value, or an interval. The response modality was formulated accordingly. It was up to the participants to give point or interval value responses. In each experimental condition the content of the cover story remained constant, the percentages were varied.

In the MODUS PONENS tasks with certain premises (100% in both premises), all participants solved the task correctly and responded "100%" ($N = 45$, [55, Experiment 1 and 2]). In the tasks with uncertain premises we observed that the participants inferred probabilities that were close to the normative values. This result mirrors the endorsement rate of 89–100% reported for the classical form of the MODUS PONENS [20]. In one experiment the participants also evaluated the negated conclusion, $\neg B$, from the same premises ($n = 30$, [55, Experiment 2]). Again, in the tasks with certain premises all participants inferred correctly "0%". These results indicate three things. First, the participants do not adopt a simple matching strategy. Second, the participants are perfect in the "certain MODUS PONENS" and the respective task with the negated conclusion. Third, the reliability of our experimental conditions is high. The results agree with the literature. Human subjects are perfectly competent to make MODUS PONENS inferences if the premises are certain.

In the MODUS PONENS tasks with uncertain premises about 70% of the responses were interval responses (averaged over different tasks). Figure 3 presents results of the MODUS PONENS [55, data from Experiment 2]. Each interval response belongs to one of the following six categories: (i) the response is coherent (the lower and the upper probabilities are both in the coherent interval), (ii) only the lower bound response is coherent, (iii) only the upper bound response is coherent, (iv) the interval response is too low, (v) the interval response is too high, and (vi) too wide interval responses. The majority of the interval responses falls into the coherent category (i).

We evaluated the agreement of the responses and the normative values by χ^2-values. The χ^2-values were calculated with the help of (i) the actual number of responses falling into the normatively correct intervals and (ii) the expected number, which was determined by the size (range) of the normative intervals (guessing assumption). High χ^2 values in the predicted direction (a high value in the opposite direction did not occur) indicate more than expected coherent responses. Compared with MODUS TOLLENS, AFFIRMING THE CONSEQUENT, and DENYING THE ANTECEDENT, the by far best

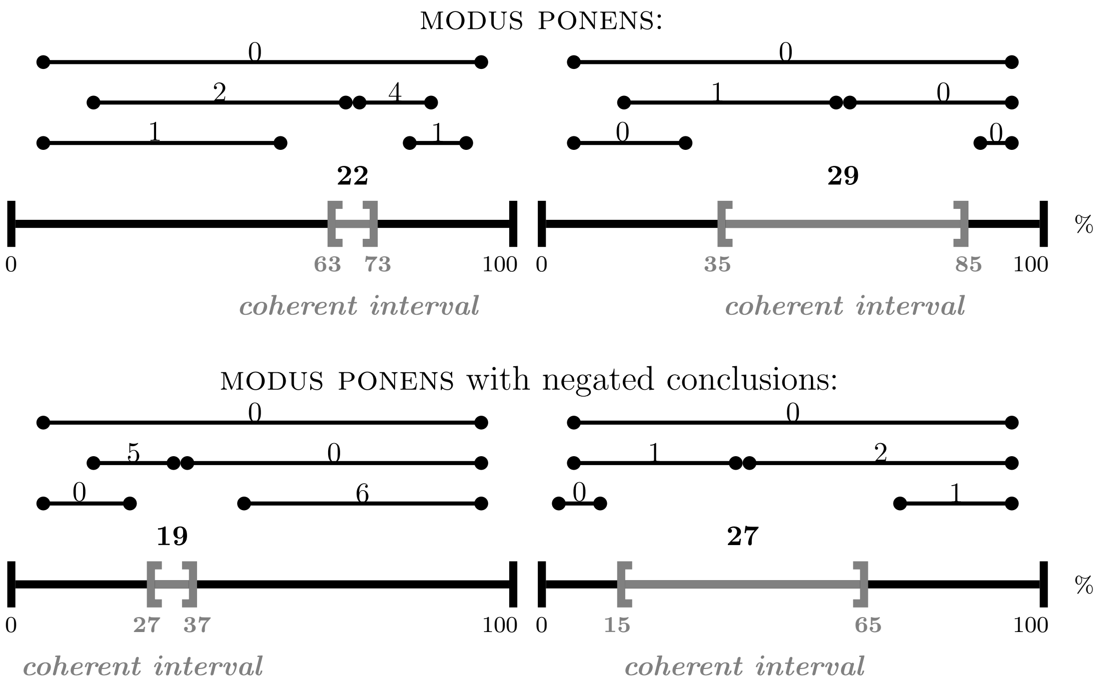

Figure 3: Frequencies of interval responses in the MODUS PONENS condition with uncertain premises ($n = 30$). In the left column the premises were $P(B|A) = .7$ and $P(A) = .9$, and in the right column the premises were $P(B|A) = .7$ and $P(A) = .5$. Each participant inferred $P(B)$ (first row) and $P(\neg B)$ (second row) from the premises. The majority of the interval responses are coherent.

agreement with the coherent intervals is observed for the MODUS PONENS [56].

To explain the difficulty of an inference task, Evans [18] proposed two important properties of the tasks, directionality and negativity. A task is forward/backward directed, if it requires inference from the antecedent/consequent to the consequent/antecedent. A task is positive/negative if it does/does not involve negation. Positive forward tasks are easy. Negative backward tasks are difficult. These are well known findings in the literature. In [56] we propose first steps for a systematic explanation of these effects. Inferring the probability of the conclusion from the premises of the backward tasks requires to build fractions, which is difficult in general. The negative tasks require the building of complements, which means further steps in the reasoning processes. The MODUS PONENS is therefore easy, because it is the most elementary positive forward task.

A third property, that may contribute to the good agreement of the actual responses and the normative lower probabilities, is the fact that normatively the lower probabilities of the conclusions are just the product of the two probabilities in the premises. The conclusion must always be considerably less than the smaller one of the two premise probabilities. The participants may have a good understanding of this fact. For first steps towards a process model based on propositional graphs see [56].

We next turn to a critical test of human probabilistic reasoning. If the probability of A is x, then the probability of its negation is the 1-complement $P(\neg A) = 1 - x$. What is the complement of a probability interval? It is given by the following equivalence

$$\boxed{P(A) \in [x', x'']} \text{ if, and only if } \boxed{P(\neg A) \in [1 - x'', 1 - x']}\ .$$

This property is called *conjugacy*. Conjugacy is a necessary condition for coherence [64]. Are humans able to infer the interval complements for negations? We observed that in the MODUS PONENS problems a surprisingly high number of lower and upper probabilities agreed perfectly with the conjugacy property [55].

4 Coherence based semantic for nonmonotonic reasoning

Nonmonotonic reasoning is a branch of artificial intelligence that, among many other branches, investigates the formalization of common sense reasoning. It investigates inferences that allow to retract conclusions in the light of new evidence. There are many systems of nonmonotonic reasoning [22]. SYSTEM P is a set of inference rules that satisfy basic rationality postulates of nonmonotonic reasoning [1, 32].[9] SYSTEM P is of particular significance,

[9]The "P" in "SYSTEM P" stands for the preferential model semantics proposed in [32].

since it is broadly accepted in the nonmonotonic reasoning community. The principles of SYSTEM P are also discussed in several other systems, weaker ones [26] and stronger ones [34, 24, 60].

The role of the conditional in SYSTEM P is of special interest in the present context. As explained above, nonmonotonic conditionals are conditionals that allow for exceptions. Nonmonotonic conditionals occur in phrases like "birds can normally fly" or just "birds can fly". Their defeasibility is often not stated explicitly. Nonmonotonic conditionals play an essential role in each inference rule of SYSTEM P. SYSTEM P determines which inferences about nonmonotonic conditionals are acceptable and which ones are not acceptable. Only weakened versions of the monotonic inferences (PREMISE STRENGTHENING, HYPOTHETICAL SYLLOGISM, etc.) are acceptable. SYSTEM P satisfies two desirable properties: (i) SYSTEM P is "weak" (or cautious) enough in the sense that the undesirable monotonicity principle does not hold, and (ii) SYSTEM P is "strong" enough to draw default conclusions, with the possibility left to withdraw them in the light of new evidence. These two properties are violated in classical logic.

If conditionals are represented by conditional events with associated probabilities, then the principles of SYSTEM P allow an interpretation in probability theory. Gilio [23] has developed a probability semantic for nonmonotonic reasoning, which is based on coherence. Coherent conditional probabilities represent nonmonotonic conditionals. The degree of normality is represented by the associated conditional probability. Gilio has shown for each rule of SYSTEM P how the (interval-)probability of the conclusion is constrained by the premises. All the rules of SYSTEM P are probabilistically informative. All the rules of SYSTEM P are p-valid [1]. Furthermore, if all probabilities in the premises are equal to 1, then the probability of the conclusion is equal to 1. If the nonmonotonic conditional is replaced by the material conditional, then the rules of SYSTEM P are logically valid.

We explain how nonmonotonic inferences are cast into a probabilistic format by the standard example of nonmonotonic reasoning, namely the Tweety problem:

> Tweety is a bird, and as you know that birds can normally fly, you conclude by default that Tweety can fly. When you learn that Tweety is a penguin, common sense tells you to retract your default conclusion that Tweety can fly.

The probabilistic version of the Tweety example runs as follows:

Premise 1:	$P[\text{Fly}(x)\vert\text{Bird}(x)] = .95.$	*(Birds can normally fly.)*
Premise 2:	Bird(Tweety).	*(Tweety is a bird.)*
Conclusion 1:	$P[\text{Fly}(\text{Tweety})] = .95.$	*(Tweety can normally fly.)*

Premise 1 and 2 are the initial premises, Conclusion 1 is the default conclusion. This inference is justified by the probabilistic MODUS PONENS. Premise

3–5 introduce new evidence. Conclusion 2 is the new (default) conclusion, after the revision in the light of the new evidence:

Premise 3:	Penguin(Tweety).	*(Tweety is a penguin.)*
Premise 4:	$P[\text{Fly}(x)\vert\text{Penguin}(x)] = .01$.	*(Penguins normally can't fly.)*
Premise 5:	$P[\text{Bird}(x)\vert\text{Penguin}(x)] = .99$.	*(Penguins are normally birds.)*
Conclusion 2:	$P[\text{Fly(Tweety)} \mid \text{Bird(Tweety)} \wedge \text{Penguin(Tweety)}] \in [0, .01]$.	
	(If Tweety is a bird and a penguin, normally Tweety can't fly.)	

The inference from Premise 1–5 to Conclusion 2 is justified by the CAUTIOUS MONOTONICITY rule of SYSTEM P (from $\boxed{A \textit{ normally } B \text{ and } A \textit{ normally } C}$ infer $\boxed{A \wedge B \textit{ normally } C}$; see [23]). Premise 4 and Premise 5 are instantiations of the premises and Conclusion 2 is an instance of the conclusion of the CAUTIOUS MONOTONICITY. CAUTIOUS MONOTONICITY is a cautious version of PREMISE STRENGTHENING. Both are discussed in the next section.

4.1 Example 2: Nonmonotonic conditionals and premise strengthening

4.1.1 Formal background

PREMISE STRENGTHENING is logically valid, since $A \supset C$ logically implies $(A \wedge B) \supset C$. For nonmonotonic [32], counterfactual [35], and causal conditionals, however, PREMISE STRENGTHENING is *not* valid. Consider the following inference involving nonmonotonic conditionals:

> IF the bird Tweety is frightened, NORMALLY Tweety flies away. Therefore: IF the bird Tweety is frightened and if Tweety is a penguin, NORMALLY Tweety flies away.

Replacing the nonmonotonic conditionals (IF–NORMALLY) by counterfactual conditionals (IF–WERE—THEN–WOULD BE), or by causal conditionals (–CAUSES THAT–), shows that PREMISE STRENGTHENING is counterintuitive. PREMISE STRENGTHENING is probabilistically not informative,

$$\textit{from } \boxed{P(C|A) = x} \textit{ infer } \boxed{P(C|A \wedge B) \in [0,1]} .$$

SYSTEM P contains a cautious version of the PREMISE STRENGTHENING, namely the CAUTIOUS MONOTONICITY, which is probabilistically informative [23],

$$\textit{from } \boxed{P(C|A) = x \textit{ and } P(B|A) = y}$$
$$\textit{infer } \boxed{P(C|A \wedge B) \in [\max\{0, (x+y-1)/y\}, \min\{x/y, 1\}]} .$$

4.1.2 Empirical investigation of monotonicity

In [49] we divided forty participants into two groups. Twenty participants received fourteen CAUTIOUS MONOTONICITY tasks and twenty participants received fourteen PREMISE STRENGTHENING tasks. The premises and the conclusions of the tasks map the corresponding inference rule (see the previous section). The content of the cover-stories was the same in both conditions. We varied the percentages in the premises within both conditions. If the participants understand the probabilistic non-informativeness of the PREMISE STRENGTHENING, they should infer wide and non-informative intervals.

In the CAUTIOUS MONOTONICITY condition, about 63% of the responses were interval responses (averaged over different tasks). In the PREMISE STRENGTHENING condition, 69% of the responses were interval responses. Figure 4 presents the frequencies of the lower and of the upper bound responses. The interval responses are clearly larger in the PREMISE STRENGTHENING condition than in the CAUTIOUS MONOTONICITY condition. In both conditions the lower and the upper bound responses are quite close to the normative values.

In the PREMISE STRENGTHENING condition, more than half of the participants responded by lower bounds $\leq 1\%$. More than half of the participants responded by upper bounds that are equal to the values presented in the premises of the tasks. On the average, 27% of the participants responded by intervals with both lower bounds $\leq 1\%$ and upper bounds $\geq 91\%$ ($n_1 = 20$, 14 tasks). Most participants understood that the lower bound can be practically zero. Most participants used a matching heuristic for inferring the upper bound. Apparently, most participants correctly inferred the lower bound of the conclusion but did not continue to search the upper bound.

If the conditional were interpreted as a material conditional, then PREMISE STRENGTHENING would be probabilistically informative [52]. Psychologically, the following prediction follows: most subjects infer $P(A \wedge B \supset C) \in [x, 1]$ from $P(A \supset C) = x$. Most participants, however, responded by lower bounds close to zero and by upper bounds close to x. Thus, most participants did not interpreted the conditional as a material conditional.

In the CAUTIOUS MONOTONICITY condition, the correlation between the mean lower bound responses and the normative lower bounds over all fourteen tasks was very high ($r = 0.92$). Subjects are sensitive to the probabilistic non-informativeness of PREMISE STRENGTHENING—at least concerning the lower bound. Subjects are cautious with monotonicity.

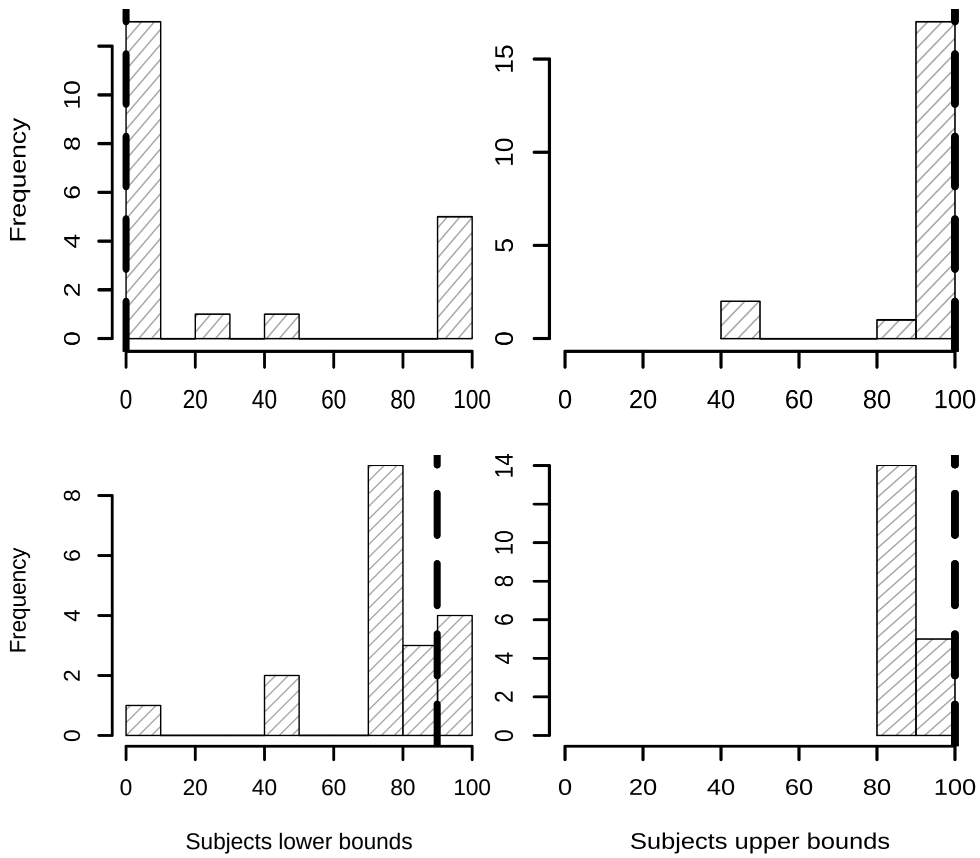

Figure 4: Frequencies of the lower and upper bound responses of Task 1 in the PREMISE STRENGTHENING condition (upper row; $n_1 = 20$) and in the CAUTIOUS MONOTONICITY condition (lower row; $n_2 = 19$) [49]. Most responses are close to the normative lower and upper bounds (dashed lines), respectively.

4.2 Example 3: Nonmonotonic conditionals and hypothetical syllogism

The HYPOTHETICAL SYLLOGISM is logically valid, since $A \supset B$ and $B \supset C$ logically imply $A \supset C$. Thus, the material conditional is transitive. The nonmonotonic conditional, however, is not transitive. The probabilistic version of the HYPOTHETICAL SYLLOGISM is probabilistically not informative,

$$\textit{from } \boxed{P(C|B) = x \textit{ and } P(B|A) = y} \textit{ infer } \boxed{P(C|A) \in [0, 1]}.$$

The HYPOTHETICAL SYLLOGISM is not contained in the nonmonotonic SYSTEM P and adding it would make SYSTEM P monotonic. However, a weaker version of the HYPOTHETICAL SYLLOGISM is contained in SYSTEM P, namely the CUT rule ("cut" is a shortcut for "cumulative transitivity"),

$$\begin{array}{ll} \textit{from} & \boxed{P(C|A \wedge B) = x \textit{ and } P(B|A) = y} \\ \textit{infer} & \boxed{P(C|A) \in [xy\,,\, xy + 1 - y]}. \end{array}$$

The CUT is a conditional version of the MODUS PONENS. The reason is clear by comparing the probability propagation rules of both argument forms and by dropping the conditioning event A.

We observed that the participants in the HYPOTHETICAL SYLLOGISM condition did not understand the probabilistic non-informativeness. They inferred informative intervals close to the coherent values of the CUT problems. We explain this result as follows [49]. Adams [1] stressed the probabilistic invalidity of the HYPOTHETICAL SYLLOGISM. He suggested to interpret its occurrence in common sense reasoning as an instantiation of CUT. Bennett [4] justified Adams' suggestion in terms of conversational implicatures [25]. If a speaker first utters a premise of the form A *normally* B and then utters as the second premise B *normally* C, the speaker actually means by the second premise a sentence of the form ($\underline{A \textit{ and}}$ B) *normally* C. The speaker does not mention "A *and*" to the addressat because A *and* is already conversationally implied and "clear" from the context. Suppose we speak (as we did in [49, Experiment 2 and 3]) about cars on a *big parking lot* that are *blue*, and suppose we then add that

> You know with 91% certainty: if a car is *blue*, then the car has *grey tire-caps* ,

you probably assume that we are speaking about the blue cars *that are on the big parking lot*, even if we do not mention this explicitly.

This interpretation explains why participants do not infer wide intervals. If the conversational implicature hypothesis is correct, then the participants actually interpret the HYPOTHETICAL SYLLOGISM tasks as instances of the CUT rule.

Another rule of SYSTEM P is the RIGHT WEAKENING rule, which is probabilistically informative and, like CUT, a cautious version of the HYPOTHETICAL SYLLOGISM,

$$\begin{array}{ll} \textit{from} & \boxed{P(B|A) = x \textit{ and } B \supset C \textit{ is logically true}} \\ \textit{infer} & \boxed{P(C|A) \in [x, 1]} \ . \end{array}$$

Practically all participants inferred correctly the lower bound in the RIGHT WEAKENING tasks [49]. The percentage of participants that inferred correctly "100%" as the upper bound varied between 50% and 75%.

4.3 Example 4: Nonmonotonic conditionals and contraposition

4.3.1 Formal background

CONTRAPOSITION belongs to the class of inferences often called "immediate" or "direct" inferences. CONTRAPOSITION is an inference from one conditional premise to a conditional conclusion. CONTRAPOSITION denotes the logical equivalence of $A \supset B$ and $\neg B \supset \neg A$. Because of this equivalence we call the corresponding argument forms "CONTRAPOSITION":

$$\textit{from } \boxed{A \supset B} \ \textit{ infer } \boxed{\neg B \supset \neg A} \ ,$$

and

$$\textit{from } \boxed{\neg B \supset \neg A} \ \textit{ infer } \boxed{A \supset B} \ .$$

If the conditional is interpreted as a material conditional, then $P(A \supset B) = P(\neg B \supset \neg A)$. Thus, the corresponding probabilistic argument form assigns exactly the same probability value to the conclusion as given in the premise. If, however, the conditional is interpreted as a conditional probability, then $P(B|A)$ is not necessarily equal to $P(\neg A|\neg B)$. The corresponding probabilistic argument forms are

$$\textit{from } \boxed{P(B|A) = x} \ \textit{ infer } \boxed{P(\neg A|\neg B) \in [0, 1]}$$

and

$$\textit{from } \boxed{P(\neg A|\neg B) = x} \ \textit{ infer } \boxed{P(B|A) \in [0, 1]} \ ,$$

for all probability values x. We see here a clear watershed between argument forms containing the material conditional and argument forms containing the conditional event. If the subjects interpret the conditional as a material conditional then a simple matching strategy is the best strategy for CONTRAPOSITION problems. If the subjects interpret the conditional as a conditional event, however, the argument is probabilistically not informative

Probabilistic CONTRAPOSITION	*Mental models* *fully explicit*	*implicit*	*Conditional probability*
$P(\text{IF } A, \text{ THEN } B) = x$	$P(A \supset B) = x$	$P(A \wedge B) = x$	$P(B\|A) = x$
$P(\text{IF } \neg B, \text{ THEN } \neg A) = ?$	$P(\neg B \supset \neg A) = x$	$P(\neg B \wedge \neg A) \in [0, 1-x]$	$P(\neg A\|\neg B) \in [0, 1]$

Table 2: Probabilistic CONTRAPOSITION in the mental model theory and in the conditional probability interpretation. The first row corresponds to the premise and the second row corresponds to the conclusion. For simplicity, the "mental footnote" is ignored. The probability of the premise is given (x), and the subject infers the probability of the conclusion. Mental model theory predicts that the premise constrains the probability of the conclusion. Mental probability logic predicts that CONTRAPOSITION is probabilistically not informative.

and any assessment of the conclusion is coherent. Thus, the material conditional interpretation predicts matching, and the conditional probability interpretation predicts non-informative probability intervals.

The mental model theory [29, 28] postulates that the core meaning of a conditional corresponds to the material conditional. Usually, human subjects represent a conditional by an implicit mental model. An implicit mental model consists of a mental model of the conjunction of the antecedent and the consequent, plus a "mental footnote" which represents the two truth table cases where the antecedent is false ($\neg A \wedge B$ and $\neg A \wedge \neg B$). All three truth table cases are just those in which $A \supset B$ is true (see Table 1). Thus, IF A, THEN B is represented by

$$\begin{array}{c} A \wedge B \\ \ldots \end{array}$$

where "..." represents the mental footnote. Because of the mental footnote, the whole representation is called the "implicit mental model" of the conditional. Under some circumstances, subjects "flesh out" the implicit mental model by replacing the mental footnote by representations of $\neg A \wedge B$ and of $\neg A \wedge \neg B$. The resulting mental model is called "fully explicit". Fully explicit mental models require much more working memory load than implicit mental models. Therefore the mental model theory claims that conditionals are represented usually by implicit and not by explicit mental models.

In general, the probability of the implicit mental model of IF A, THEN B is equal to $P(A \wedge B) + P(\ldots)$. We assume that the probabilistic assessment of the truth table cases is coherent. The mental model theory assumes that the subjects focus on $P(A \wedge B)$. For simplicity, we assume that the probability of the "mental footnote", $P(\ldots)$, is ignored by the subject. Thus, we make no special assumptions about $P(\ldots)$. If the subjects form implicit mental models, and if $P(\ldots)$ is ignored, then the CONTRAPOSITION inference consists of inferring $P(\neg A \wedge \neg B)$ from $P(A \wedge B)$. If $P(A \wedge B) = x$, then

$P(\neg A \wedge \neg B) \in [0, 1-x]$, because the truth table cases must add up to one.

The probability of the fully explicit mental model is equal to $P(A \wedge B) + P(\neg A \wedge B) + P(\neg A \wedge \neg B)$. The probability of the fully explicit mental model corresponds to the probability of the material conditional. Table 2 presents the implicit and the fully explicit mental models of the CONTRAPOSITION. CONTRAPOSITION is probabilistically informative in the mental model theory. CONTRAPOSITION is probabilistically not informative in the mental probability logic (because of the conditional probability interpretation of the IF–THEN). The next section investigates these claims empirically.

4.3.2 Empirical results of the contraposition

We presented (among other tasks) both forms of CONTRAPOSITION to the participants [53]. In the CONTRAPOSITION task with the negations in the premise, 58% of the participants inferred both lower bounds $\leq$ 7% and upper bounds $\geq$ 93% ($n_1 = 40$). These wide interval responses indicate understanding of the probabilistic non-informativeness of the argument. In the CONTRAPOSITION task with the negations in the conclusion, the respective percentage was 40% ($n_2 = 40$).

Practically all of the participants who did not infer such wide intervals responded either by point values close to zero, or by point values that are close to one hundred. We believe that these participants stop to infer probabilities as soon as they see one extreme bound of the interval of the conclusion and take this as the probability of the conclusion. Apparently, they understand that the one extreme pole of the unit interval is a coherent assessment of the conclusion and neglect to continue to search for the other one.

In the CONTRAPOSITION task with the negations in the premise, only three of the forty participants reproduced the value presented in the premise. In the CONTRAPOSITION task with the negations in the conclusion, ten of the forty participants reproduced the value presented in the premise. Over all conditions only thirteen of the eighty participants responded by reproducing the percentage contained in the premise. Thus only 16% of the participants conform with the probability of the material conditional interpretation (fully explicit mental model). As noted above, it is not clear whether these participants actually interpret the conditional as a material conditional or whether they simply apply a matching strategy. Nevertheless, this is strong evidence against the material conditional interpretation.

In the CONTRAPOSITION task with the negations in the premise, only four of the forty participants responded by the $[0\%, 7\%]$ interval, which corresponds to the implicit mental model. In the CONTRAPOSITION task with the negations in the conclusion, none of the responses corresponds to the implicit mental model. Overall, only 10% of the interval responses of the participants correspond to the implicit mental model.

The results challenge two predictions of the mental model theory. First, the broad majority of the interval responses should be explained by implicit and explicit mental models. Second, more implicit mental models should be observed than explicit mental models. Our data do not support these predictions.

Finally as a technical remark, CONTRAPOSITION implies the monotonicity property of classical logic. Since most participants understand that CONTRAPOSITION is probabilistically not informative their reasoning can be interpreted as nonmonotonically and not as monotonic.

5 Concluding remarks

We described a probabilistic approach to human conditional reasoning. Nonmonotonic conditionals were investigated in the framework of probability logic. We selected a special probability logic which is based on coherence. It combines logic and subjective probability theory. Probability logic tells us how to infer deductively the coherent (lower and upper) probability of the conclusion from the premises. The structure of the inference task is analyzed in a rigorous way: everything known is made explicit in the premises. We make only one implicit assumption, namely that the assessment of the premises must be coherent. The cover-stories of our tasks were designed to map the structure of the probability logical arguments. We took special care that the content of the cover-stories is neutral and that they don't evoke uncontrolled background knowledge in the subjects.

In some recent studies on the conditional, the experimenter provided the subjects the probabilities of *all* possible states (rows in the truth table, elementary events, constituents) in the sample space. This corresponds to a complete knowledge of the joint probability distribution. Logical argument forms, though, contain only a few premises, and their probabilistic versions contain correspondingly only a few probabilities (marginal and conditional probabilities). When the probabilities of all possible states are known, the inference task is substantially different from its original logical argument form. As an example, assume the following assessment of all possible states in the sample space:

$$P(A \wedge B) = .5,\ P(A \wedge \neg B) = .1,\ P(\neg A \wedge B) = .2,\ P(\neg A \wedge \neg B) = .2$$

This assessment describes a situation of complete probabilistic knowledge. In this situation, the premises of the MODUS PONENS are $P(A) = .60$ and $P(B|A) = .83$, and the conclusion is $P(B) = .70$. If, however, the truth table cases are unknown, and all we know is that $P(A) = .60$ and that $P(B|A) = .83$, then we are in a situation of incomplete probabilistic knowledge. Then we get an interval in the conclusion, $P(B) \in [.50, .90]$ (see Figure 2). This example shows the necessity to work with interval probabilities.

Since situations of complete probabilistic knowledge are seldom in everyday life, experimenters should also focus on reasoning from incomplete probabilistic knowledge. Interval probabilities are psychologically highly plausible and we presented empirical evidence that they are actually used when offered as a response mode. We observed that subjects are especially good in inferring the lower probability bound of the conclusion of the probabilistic MODUS PONENS.

The four conditional inferences (MODUS PONENS, etc.) are prominent in psychology and were investigated several times. We investigated argument forms that go beyond those four conditional inferences. PREMISE STRENGTHENING, CONTRAPOSITION, and HYPOTHETICAL SYLLOGISM and the rules of SYSTEM P are conditional argument forms. They involve more general properties of conditionals. The material conditional is transitive and monotonic. The nonmonotonic conditional is neither transitive nor monotonic.

We observed that many subjects understand the probabilistic non-informativeness of PREMISE STRENGTHENING and of CONTRAPOSITION. The interval responses in the corresponding SYSTEM P rules were close to the normative values. Subjects seem not to understand the probabilistic non-informativeness of HYPOTHETICAL SYLLOGISM. We explained this result by conversational implicatures. Subjects interpret the HYPOTHETICAL SYLLOGISM problems as instances of the CUT rule, which is a corresponding SYSTEM P rule. For some inference rules of SYSTEM P on may speculate that they are at the core of the human inference engine (especially the LEFT LOGICAL EQUIVALENCE and the RIGHT WEAKENING rules, see [50, 48]).

In our experiments, we presented the uncertainty of the premises in terms of percentages. We wanted to avoid verbal paraphrases. Phrases like "probably" are ambiguous. The subject must infer the meaning of such phrases. Inferences to an interpretation of verbal paraphrases bias the experiment, since different participants may infer different meanings. By presenting percentages in the premises, the degree of uncertainty is controlled in the experiment. The investigation of reasoning *to an interpretation* is important for the understanding of how subjects form representations. In our studies, however, we were concerned with reasoning *from an interpretation*, which refers to how subjects manipulate representations from fixed interpretations. The importance of the distinction between reasoning to an interpretation and reasoning from an interpretation is stressed by Stenning and van Lambalgen [61, 62]. Future work will investigate cognitive representations and processing of the probabilities.

While we were primarily concerned with reasoning from fixed premises, our studies also provide insight into how humans interpret the IF–THEN. The CONTRAPOSITION problem, for example, provides a clear watershed between the probability of the material conditional interpretation and the conditional probability interpretation. Our data clearly favor the conditional probability interpretation of the IF–THEN. This result adds to the results of Over, Evans

& Elqayam [45]. Moreover, we showed how predictions of the probabilistic extension of the mental model theory can be expressed in the language of probability logic. These predictions were not corroborated.

We investigated probabilistic inference problems that mirror central properties of nonmonotonic reasoning and properties of monotone logics. We have not yet investigated withdrawing conclusions in the light of new evidence, but our data corroborate broadly some basic rationality postulates of nonmonotonic reasoning. Future empirical work will be devoted to weaker and stronger systems than SYSTEM P. Moreover, we plan to use computer controlled experiments, take reaction times, and investigate the revision of default conclusions.

Acknowledgments

We thank David Over and the referee for helpful comments. This work is supported by the Austrian Research Fonds, FWF (project P20209 "Mental probability logic").

References

[1] E. W. Adams. *The logic of conditionals.* Reidel, Dordrecht, 1975.

[2] E. W. Adams. *A Primer of Probability Logic.* CSLI, Stanford, 1998.

[3] S. Benferhat, J.-F. Bonnefon, and R. Da Silva Neves. An overview of possibilistic handling of default reasoning, with experimental studies. *Synthese*, 1-2:53–70, 2005.

[4] J. Bennett. *A philosophical guide to conditionals.* Oxford University Press, Oxford, 2003.

[5] J.-F. Bonnefon. Title of chapter. Chapter XY, this book., 2008.

[6] J.-F. Bonnefon and D. J. Hilton. The suppression of modus ponens as a case of pragmatic preconditional reasoning. *Thinking & Reasoning*, 8(1):21–40, 2002.

[7] J.-F. Bonnefon and D. J. Hilton. Consequential conditionals: Invited and suppressed inferences from valid outcomes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1):28–37, 2004.

[8] M. D. S. Braine and D. P. O'Brien, editors. *Mental logic.* Erlbaum, Mahwah, 1998.

[9] R. M. J. Byrne. Suppressing valid inferences with conditionals. *Cognition*, 31:61–83, 1989.

[10] R. M. J. Byrne, O. Espino, and C. Santamaría. Counterexamples and the suppression of inferences. *Journal of Memory and Language*, 40:347–373, 1999.

[11] N. Chater and M. Oaksford. The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38:191–258, 1999.

[12] N. Chomsky. *Aspects of the theory of syntax*. MIT Press, Cambridge, 1965.

[13] G. Coletti and R. Scozzafava. *Probabilistic logic in a coherent setting.* Kluwer, Dordrecht, 2002.

[14] B. De Finetti. Foresight: Its logical laws, its subjective sources (1937). In H. Jr. Kyburg and H. E. Smokler, editors, *Studies in subjective probability*, pages 55–118. Wiley, New York, 1964.

[15] B. De Finetti. *Theory of probability*, volume 1, 2. John Wiley & Sons, Chichester, 1974. Original work published 1970.

[16] K. Dieussaert, W. De Neys, and W. Schaeken. Suppression and belief revision, two sides of the same coin? *Psychologica Belgica*, 45(1):29–46, 2005.

[17] G. J. W. Dorn. Popper's law of the excess of the probability of the conditional over the conditional probability. *Conceptus*, 26(67):6–61, 1992.

[18] J. S. B. T. Evans. *The psychology of deductive reasoning*. Routledge, London, 1982.

[19] J. S. B. T. Evans, S. H. Handley, and D. E. Over. Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29:321–355, 2003.

[20] J. S. B. T. Evans, S. E. Newstead, and R. M. J. Byrne. *Human Reasoning*. Erlbaum, Hove, 1993.

[21] J. S. B. T. Evans, D. E. Over, and S. J. Handley. Suppositions, extensionality, and conditionals: A critique of the mental model theory of Johnson-Laird and Byrne (2002). *Psychological Review*, 112(4):1040–1052, 2005.

[22] D. M. Gabbay and C. J. Hogger, editors. *Handbook of logic in artificial intelligence and logic programming*, volume 3. Nonmonotonic reasoning and uncertain reasoning. Clarendon Press, Oxford, 1994.

[23] A. Gilio. Probabilistic reasoning under coherence in System P. *Annals of Mathematics and Artificial Intelligence*, 34:5–34, 2002.

[24] M. Goldszmidt and J. Pearl. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence*, 84:57–112, 1996.

[25] H. P. Grice, editor. *Studies in the way of words.* Harvard University Press, Cambridge, 1989.

[26] J. Hawthorne and D. Makinson. The quantitative/qualitative watershed for rules of uncertain inference. *Studia Logica*, 86:1–52, 2007.

[27] P. N. Johnson-Laird. *Mental models: Towards a cognitive science of language, inference and consciousness.* Cambridge University Press, Cambridge, 1983.

[28] P. N. Johnson-Laird and R. M. J. Byrne. Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109(4):646–678, 2002.

[29] P. N. Johnson-Laird, V. Girotto, P. Legrenzi, and M. S. Legrenzi. Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, 106(1):62–88, 1999.

[30] P. N. Johnson-Laird and E. Shafir, editors. Reasoning and Decision Making. *Cognition, Special Issue*, 49, 1993.

[31] W. Kneale and M. Kneale. *The development of logic.* Clarendon Press, Oxford, 1984.

[32] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.

[33] F. Lad. *Operational subjective statistical methods: A mathematical, philosophical, and historical introduction.* Wiley, New York, 1996.

[34] D. Lehmann and M. Magidor. What does a conditional knowledgebase entail? *Artificial Intelligence*, 55(1):1–60, 1992.

[35] D. Lewis. Counterfactuals and comparative possibility. In W. L. Harper, R. Stalnaker, and G. Pearce, editors, *Ifs*, pages 57–85. Reidel, Dordrecht, 1973.

[36] I.-M. Liu. Conditional reasoning and conditionalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(4):694–709, 2003.

[37] I.-M. Liu, K.-C. Lo, and J.-T. Wu. A probabilistic interpretation of 'If—Then'. *The Quarterly Journal of Experimental Psychology*, 49(A):828–844, 1996.

[38] J. Macnamara. *The place of logic in psychology.* MIT Press, Cambridge, 1986.

[39] D. Marr. *Vision. A computational investigation into the human representation and processing of visual information.* W. H. Freeman, San Francisco, 1982.

[40] M. Oaksford and N. Chater. Against logicist cognitive science. *Mind & Language*, 6(1):1–38, 1991.

[41] M. Oaksford and N. Chater. A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101:608–631, 1994.

[42] M. Oaksford and N. Chater. *Bayesian rationality: The probabilistic approach to human reasoning.* Oxford University Press, Oxford, 2007.

[43] M. Oaksford, N. Chater, and J. Larkin. Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26:883–899, 2000.

[44] K. Oberauer and O. Wilhelm. The meaning(s) of conditionals: Conditional probabilities, mental models and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29:680–693, 2003.

[45] D. Over, J. S. T. B. Evans, and S. Elqayam. Conditionals and non-constructive reasoning. Chapter XY, this book., 2008.

[46] D. E. Over and J. S. B. T. Evans. The probability of conditionals: The psychological evidence. *Mind & Language*, 18:340–358, 2003.

[47] D. E. Over, C. Hadjichristidis, J. S. Evans, S. J. Handley, and S. A. Sloman. The probability of causal conditionals. *Cognitive Psychology*, 54:62–97, 2007.

[48] N. Pfeifer. Contemporary syllogistics: Comparative and quantitative syllogisms. In G. Kreuzbauer and G. J. W. Dorn, editors, *Argumentation in Theorie und Praxis: Philosophie und Didaktik des Argumentierens*, pages 57–71. LIT, Wien, 2006.

[49] N. Pfeifer and G. D. Kleiter. Nonmonotonicity and human probabilistic reasoning. In *Proceedings of the 6^{th} Workshop on Uncertainty Processing*, pages 221–234, Hejnice, 2003. September 24–27^{th}, 2003.

[50] N. Pfeifer and G. D. Kleiter. Coherence and nonmonotonicity in human reasoning. *Synthese*, 146(1-2):93–109, 2005.

[51] N. Pfeifer and G. D. Kleiter. Towards a mental probability logic. *Psychologica Belgica*, 45(1):71–99, 2005. Updated version at: `http://users.sbg.ac.at/~pfeifern/`.

[52] N. Pfeifer and G. D. Kleiter. Inference in conditional probability logic. *Kybernetika*, 42:391–404, 2006.

[53] N. Pfeifer and G. D. Kleiter. Is human reasoning about nonmonotonic conditionals probabilistically coherent? In *Proceedings of the 7th Workshop on Uncertainty Processing*, pages 138–150, Mikulov, 2006. September 16–20$^{\text{th}}$, 2006.

[54] N. Pfeifer and G. D. Kleiter. Towards a probability logic based on statistical reasoning. In *Proceedings of the 11th IPMU Conference (Information Processing and Management of Uncertainty in Knowledge-Based Systems)*, pages 2308–2315, Paris, 2006. Edition E.D.K.

[55] N. Pfeifer and G. D. Kleiter. Human reasoning with imprecise probabilities: Modus ponens and Denying the antecedent. In *5th International Symposium on Imprecise Probability: Theories and Applications*, pages 347–356, Prague, Czech Republic, 16 - 19 July 2007.

[56] N. Pfeifer and G. D. Kleiter. Framing human inference by coherence based probability logic. *Journal of Applied Logic*, in press.

[57] G. Politzer. Uncertainty and the suppression of inferences. *Thinking & Reasoning*, 11(1):5–33, 2005.

[58] F. P. Ramsey. Truth and probability (1926). In D. H. Mellor, editor, *Foundations. Essays in philosophy, logic, mathematics and economics*, pages 58–100. Routledge & Kegan Paul, London, 1978.

[59] L. J. Rips. *The psychology of proof: Deductive reasoning in human thinking.* MIT Press, Cambridge, 1994.

[60] G. Schurz. Probabilistic default reasoning based on relevance and irrelevance assumptions. In D. Gabbay et al., editor, *Qualitative and Quantitative Practical Reasoning*, number 1244 in LNAI, pages 536–553. Springer, Berlin, 1997.

[61] K. Stenning and M. van Lambalgen. A little logic goes a long way: Basing experiment on semantic theory in the cognitive science of conditional reasoning. *Cognitive Science*, 28:481–529, 2004.

[62] K. Stenning and M. van Lambalgen. The logical response to a noisy world. Chapter XY, this book., 2008.

[63] G. Störring. Experimentelle Untersuchungen zu einfachen Schlußprozessen. *Archiv für die Gesamte Psychologie*, 11:1–127, 1908.

[64] P. Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, London, 1991.

Available online at www.sciencedirect.com
ScienceDirect
Journal of Applied Logic 7 (2009) 206–217
JOURNAL OF APPLIED LOGIC
www.elsevier.com/locate/jal

Framing human inference by coherence based probability logic ☆

Niki Pfeifer *, Gernot D. Kleiter

Department of Psychology, University of Salzburg, Austria

Received 4 July 2007; received in revised form 30 October 2007; accepted 7 November 2007

Available online 28 December 2007

Abstract

We take coherence based probability logic as the basic reference theory to model human deductive reasoning. The conditional and probabilistic argument forms are explored. We give a brief overview of recent developments of combining logic and probability in psychology. A study on conditional inferences illustrates our approach. First steps towards a process model of conditional inferences conclude the paper.

Keywords: Human reasoning; Conditional; Probability logic; Argument forms

To empirically investigate human deductive inference one needs a description of what deductive inference is all about. Such a description specifies what the human mind should compute, which conclusions should be considered rational and which ones not. From Aristotle until the end of the twentieth century, classical logic was the standard reference in psychology. The emerging logical pluralism and the many new paradigms developed in computer science cast doubts upon the general appropriateness of classical logic as the standard frame in the psychology of thinking and reasoning. Recently, a strong trend in psychology emerged to consider probability to be relevant even in tasks in which uncertainties are not explicitly mentioned.

The present contribution takes probability logic based on the *coherence* approach of subjective probability as the basic reference theory. It gives a brief overview of the recent developments of combining logic and probability to build normative and descriptive models of human deductive reasoning. It explains the reasons why we think that the coherence approach offers advantages for psychological model building. We also describe results of our own experimental studies.

Coherence is a key concept in subjective probability theory. In the betting interpretation, coherence guarantees the avoidance of sure losses (often called a "Dutch book"). From a psychological perspective, the coherence approach provides several advantages. Most importantly, the coherence approach is based on the subjective interpretation of probabilities. Subjective probabilities are degrees of belief and are conceived as coherent descriptions of incomplete knowledge states. While human reasoning may be more or less coherent, it in any case involves degrees of belief and descriptions of incomplete knowledge states. It would be an unwise research strategy to take a reference theory that is

☆ Supported by the Austrian Recearch Fonds, FWF (project P20209, Mental probability logic).
* Corresponding author.
E-mail addresses: niki.pfeifer@sbg.ac.at (N. Pfeifer), gernot.kleiter@sbg.ac.at (G.D. Kleiter).

1570-8683/$ – see front matter
doi:10.1016/j.jal.2007.11.005

far away from cognitive science, for instance, an approach to probability developed for thermodynamics or quantum physics.

The coherence approach to probability offers conceptions that are not available in other approaches. Examples are the renouncement of full algebras of events, the special conception of conditional probability, the treatment of single events, the expression of imprecise probabilities, the closeness to Bayesian statistics, and, last but not least, the availability of probability semantics for logical systems such as SYSTEM P.

There are at least two scientific communities in which *coherence* is basic. The "Italian community" and the "imprecise probability community". De Finetti [9] is the root of both. Much of the recent work on the coherence approach is summarized in [6,22]. The standard reference of imprecise probability is [38]. In recent years the Italian community has connected its approach to possibility theory and other systems of uncertainty representation. The imprecise approach has been extended to statistical modeling, conditional independence models, and several other fields of uncertainty processing.

1. Conditional

Recently, several psychologists have followed the proposal of Adams [1,2], Popper [34], Rényi [37], and others and interpret the indicative conditional IF H, THEN E as a conditional event, $E|H$ (with the according probability $P(E|H)$), and not as a material conditional, $H \supset E$ (with the according probability $P(H \supset E)$). The "mental model" theory may be an exception [5,19].

Interpreting the probability of the IF–THEN as the probability of the material implication leads to the paradoxes of the material implication. This is not the case if the IF–THEN is interpreted as a conditional event. For instance, the paradoxical argument form $\neg H \therefore$ IF H, THEN E can be probabilistically interpreted as $P(\neg H) = x \therefore P(H \supset E) \in [x, 1]$, where $0 \leqslant x \leqslant 1$. Here the conclusion is constrained by the premise. If the IF–THEN is interpreted as a conditional event, however, only $P(E|H) \in [0, 1]$ follows. The conclusion is not constrained by the premise. This is an advantage of the conditional probability interpretation, since the premises of a counterintuitive argument form should provide no information about its conclusion. Moreover, the probability values of $P(H \supset E)$ and $P(E|H)$ can differ substantially. Dorn [10] considers a compelling example. Let H be THE NEXT THROW OF THE DIE WILL COME UP A 5, and let E be THE NEXT THROW OF THE DIE WILL COME UP AN EVEN NUMBER. For a fair die $P(H) = 1/6$ and $P(E) = 1/2$. Then, $P(H \supset E) = 5/6$, whereas $P(E|H) = 0$.

Introducing a conditional as a conditional event is special in the coherence approach. First, while usually—for example in the Kolmogorov approach—conditional probabilities are *defined* by absolute probabilities, in the coherence approach conditional events are basic. "In classical approaches, a conditional probability $P(E|H)$ is not introduced as a *direct* notion, and so there is no meaning given to $E|H$ itself" [8]. If conditional probabilities are "defined" by absolute probabilities, then there is no conditional entity *per se*, and also no non-probabilistic IF–THEN. A conditional event is not an ordered pair (E, H), with $H \neq \emptyset$, where $\emptyset$ is the impossible event [6, p. 63]. Assigning probability values directly to the conditional event is psychologically highly plausible: a person does not need to know and process the "joint" and "marginal" probabilities to come up with a conditional probability assessment (where "conditional = joint/marginal").

Second, as in the coherence approach conditional events are basic, we may reflect upon the behavior of their truth values. They require special attention. What are the truth values of a conditional event $E|H$? If H is true the answer is straightforward,

$$|E \mid H| = \begin{cases} 1 & \text{if } E = 1 \text{ and } H = 1 \\ 0 & \text{if } E = 0 \text{ and } H = 1. \end{cases}$$

Here 0 and 1 are the indicator values corresponding to the truth values FALSE and TRUE. What is the truth value of $E|H$ when H is false? For this case de Finetti proposed a third truth value "undetermined". A similar proposal was made by Ramsey in 1929 [36, footnote, p. 155].[1]

[1] In 1926 Ramsey already introduced "the degree of belief in p given q" [35, p. 82]. He noted that "this does not mean the degree of belief in 'If p then q' [material implication], or that in 'p entails q' ". This is a misprint we found in several reprints of the article; "p" should be replaced by "q", and vice versa.

"If two people are arguing 'If p will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q; ... We can say they are fixing their degrees of belief in q given p. If p turns out false, these degrees of belief are rendered *void*."

In terms of bets, one neither wins nor loses if the conditioning event turns out to be false. The bet is annulled and the ticket prize is payed back. A consequence of such a conception is that conditioning cannot be expressed by operators like negation, conjunction, and disjunction. There is no logical operator of conditioning [17]. This is a fundamental property that distinguishes | from $\supset$. The vertical stroke operator, |, is not truth-functional, while the material implication, $\supset$, is truth-functional.

When conditional events are considered to be primitive, then probability axioms should be introduced for conditional events. An elegant method to introduce and justify axioms of conditional events was proposed by Coletti and Scozzafava [6,8]. We first give the axioms and then their justification.

Definition 1 *(Conditional probability).* Let $\mathcal{C} = \mathcal{G} \times \mathcal{B}^0$ be a set of conditional events $\{E|H\}$ such that $\mathcal{G}$ is a Boolean algebra and $\mathcal{B} \subseteq \mathcal{G}$ is closed with respect to (finite) logical sums, with $\mathcal{B}^0 = \mathcal{B} \setminus \{\emptyset\}$. A function $P : \mathcal{C} \mapsto [0, 1]$ is a conditional probability iff the following three axioms are satisfied

A1 $P(H|H) = 1$, for every $H \in \mathcal{B}^0$, or (equivalently) $P(E|H) = P(E \wedge H|H)$,
A2 $P(\cdot|H)$ is a (finitely additive) probability on $\mathcal{G}$ for any given $H \in \mathcal{B}^0$,
A3 $P(E \wedge A|H) = P(E|H)P(A|E \wedge H)$ for any $A, E \in \mathcal{G}, H, E \wedge H \in \mathcal{B}^0$.

In the present context we interpret the "$E|H$" as IF H, THEN E. Axiom A2 specifies that, for a fixed antecedent, the probabilities of the consequents follow the rules of finite additive probability. Axiom A3 is a "multiplicative chain" rule for conjunctions.

To define the "truth-value" $T(E|H)$ of a conditional event $E|H$, Coletti and Scozzafava consider the function

$$T(E|H) = \left\{ \begin{array}{ll} 1 & \text{if } E \wedge H \\ 0 & \text{if } \neg E \wedge H \\ p(E|H) & \text{if } \neg H \end{array} \right\} = 1 \cdot I_{E \wedge H} + 0 \cdot I_{\neg E \wedge H} + p(E|H) \cdot I_{\neg H}. \tag{1}$$

I denotes the indicator of the according event. The three values 1, 0, and $p(E|H)$ correspond to "win", "lose", "get money back", respectively. The sum of the three terms is thus a random variable,

$$X = \sum_{k=1}^{3} x_k I_{E_k}, \tag{2}$$

where $x_1 = 1, x_2 = 0, x_3 = p(E|H)$, and $E_1 = E \wedge H$, $E_2 = \neg E \wedge H$, and $E_3 = \neg H$. Let X and Y be two such three-term random variables. In the same way as two events can be combined by a logical operator to obtain a third event, say, $A \oplus B = C$, in the same way the two random variables, say X and Y, can be combined by a numerical operator $+$ to obtain a third random variable, $X + Y = Z$. What happens when two three-term random variables are added to obtain a third one? When done without specific constraints the result does not remain in the family of numbers that represents the set of conditional events. The operation may lead to a number that does not represent any event. When we consider, however, only conditional events with a fixed conditioning event and when we further consider only mutually exclusive conditioned events, $E \wedge A = \emptyset$ (so that also $E \wedge A \wedge H = \emptyset$), the function (1) is additive. In the domain of conditional events this corresponds to the disjunction of $E|H$ and $A|H$. The disjunction operator "$\vee$" for E and A given H thus corresponds to the addition operator "$+$" for the corresponding random variables. The same can be done for the conjunction operator "$\wedge$" and the (then corresponding) multiplication operator "$\times$", for events such as in axiom A3.

There is a connection between coherent conditional probability and possibility distributions [4,7]. A possibility distribution may be conceived as a standardized likelihood. The likelihood is a function of the *conditioning* events, $P(E|H_i)$, where H_i is the variable. Likelihoods are not probabilities. Operators on possibilities typically involve maxima for disjunctions, $\Pi(A \vee B) = \max\{\Pi(A), \Pi(B)\}$, and t-norms for conjunctions. Sometimes human subjects seem to confuse the "direction" of conditioning (e.g., in the well known Linda task). Standardized likelihoods might be candidates to model such cases in a "rational" way. Moreover, human max or min responses are hard to distinguish from superficial "matching" responses. In psychology "matching" means to restate numbers or other material already contained in the description of a problem, as the "solution" of the problem.

2. Properties of probabilistic argument forms

The lower and upper probabilities of elementary arguments are obtained by the method of cases together with some algebra, those of complex arguments by linear programming. An alternative method is used in the "Check Coherence" software [3]. Here are two examples of elementary arguments.

Example 1 *(MODUS PONENS).* The non-probabilistic MODUS PONENS infers B from the set of premises $\{(\text{IF } A, \text{THEN } B), A\}$. In the probabilistic version, the probabilities $P(B|A) = y$ and $P(A) = x$ are given, $P(B)$ is sought. By the theorem of total probability we have $P(B) = P(A)P(B|A) + P(\neg A)P(B|\neg A)$. The lower probability of B, $P(B) = z'$, is obtained by (case 1) assuming $P(B|\neg A) = 0$, so that $z' = xy$. The upper probability is obtained by (case 2) assuming $P(B|\neg A) = 1$, so that $z'' = xy + (1 - x)1 = 1 - x(1 - y)$.

Example 2 *(MODUS TOLLENS).* The non-probabilistic MODUS TOLLENS infers $\neg A$ from $\{(\text{IF } A, \text{THEN } B), \neg B\}$. Let $P(A) = x$, $P(B|A) = y$, $P(B) = z$, and $P(B|\neg A) = q$; in the probabilistic version y and $1 - z$ are given, $(1 - x)'$ and $(1 - x)''$ are sought. By the theorem of total probability we have $z = xy + (1 - x)q$. Solving for $1 - x$ we get $1 - x = 1 - (z - q)/(y - q) = (y - z)/(y - q)$. We distinguish three cases. (a) $q = z$ leads to the upper probability $(1 - x)'' = 1$, (b) $q = 0$ leads to the lower probability $(1 - x)' = 1 - z/y$, and (c) $q = 1$ leads to the lower probability $(1 - x)' = (z - y)/(1 - y)$. The lower probability is thus $\max\{1 - z/y, (z - y)/(1 - y)\}$.

The numerical solutions of complex problems may be found by linear programming. Let us consider an inference problem with n variables and m premises. The probability vector of the premises is denoted by $\mathbf{p} = (p_1, p_2, \ldots, p_m)$. We build a coefficient matrix $\mathbf{V}$ with $m + 1$ rows, one row for each premise and one additional row containing 1s only. Each column is associated with one of the combinatorially possible 0/1 patterns of the n variables. In the case of logical independence there are $r = 2^n$ such patterns. In the case of logical dependence there are fewer cases (or constituents; for how to obtain the constituents see, e.g., [14,15]). The v_{ij} values, $i = 1, \ldots, m$ and $j = 1, \ldots, r$, are equal to the (generalized) indicator values that premise i obtains under the truth values of the constituent j. The values are either 0, 1, or a conditional probability (for conditional events with negated conditioning events). The values are determined according to Eq. (1).

The matrix $\mathbf{V}$ together with the vector $\mathbf{p}$ builds a system of $m + 1$ linear equations with r unknowns.

$$\begin{aligned}
v_{11}\pi_1 + v_{12}\pi_2 &+ \cdots + v_{1r}\pi_r &= p_1 \\
v_{21}\pi_1 + v_{22}\pi_2 &+ \cdots + v_{2r}\pi_r &= p_2 \\
\cdots + \cdots &+ \cdots + \cdots &= \cdots \\
v_{m1}\pi_1 + v_{m2}\pi_2 &+ \cdots + v_{mr}\pi_r &= p_m \\
\pi_1 + \pi_2 &+ \cdots + \pi_r &= 1
\end{aligned} \tag{3}$$

The $(\pi_1, \pi_2, \ldots, \pi_r)$ are the unknown probabilities of the constituents. The sum of these probabilities is 1. If the number of premises is less than $r - 1$, then the linear system has no exact solution. Next we introduce the conclusion. It is represented by the objective function

$$w_1\pi_1 + w_2\pi_2 + \cdots + w_r\pi_r, \tag{4}$$

where the coefficients $w_1, w_2, \ldots, w_r$ are determined by Eq. (1).

Example 3 *(MODUS TOLLENS).* The upper part of Table 1 shows the four constituents for two variables. The lower part gives the two premises of the MODUS TOLLENS and the coefficients of the linear system. The lower and upper probabilities of the conclusion $\neg A$ with the coefficients $(1, 1, 0, 0)$ are the minimum and maximum, respectively, of the objective function $p_3 = 1 \cdot \pi_1 + 1 \cdot \pi_2$.

Table 1
MODUS TOLLENS. The upper part contains the constituents, the lower part the coefficient matrix **V**

A	0	0	1	1	
B	0	1	0	1	
$B\|A$	p_1	p_1	0	1	p_1
$\neg B$	1	0	1	0	p_2
	1	1	1	1	1

With the help of linear programming the lower and the upper values of the function are determined. If the probabilities of the premises are given in the form of intervals only, then the lower and upper probabilities of the conclusion are found by fractional programming, which requires several linear programming steps in succession [22].

An important property of arguments is the presence or absence of various kinds of logical and functional dependencies in sets of events. We first consider unconditional events. Logical independence is defined as follows.

Definition 2 *(Logical independence).* Let $\{E_1, \ldots, E_m\}$ be a set of m unconditional events. If all 2^m atoms are possible conjunctions, then the set of events is logically independent. Otherwise they are dependent.

We note that logical independence and dependence refer to a set of events. The mutual independence of two events is a special case.

We next consider the case of conditional events. To define logical independence for conditional events we follow [15].

Definition 3 *(Logical independence of conditional events).* A set of m conditional events is logically independent, if the number of constituents is 3^m.

The constituents are constructed by the combinations of the $(E_i \wedge H_i) \vee (\neg E_i \wedge H_i) \vee \neg H_i$, $i = 1, \ldots, m$. For details we refer to [14,15].

We next consider linear dependence/independence. Let $\mathbf{V}_{m+2}$ be the coefficient matrix of the premises together with the conclusion.

Theorem 1 *(Linear dependence). If the rank $r(\mathbf{V}_m + 1) = k$ and the rank $r(\mathbf{V}_{m+2}) = k + 1$, then the premises and the conclusion are linearly independent. If $r(\mathbf{V}_m + 1) = r(\mathbf{V}_{m+2})$, then the conclusion is linearly dependent on the premises.*

The rank determines the number of dimensions of a linear space.

Closely related to the dependence/independence properties is de Finetti's Fundamental Theorem [9, p. 112].

Theorem 2 *(Fundamental Theorem). Given the probabilities $P(E_1), P(E_2), \ldots, P(E_m)$ of a finite number of events, the probability of a further event E_{m+1},*

$$P(E_{m+1}) \text{ is } \begin{cases} \text{precise} & \text{if } E_{m+1} \text{ is linearly dependent on } \{E_1, E_2, \ldots, E_m\}, \\ \in [0,1] & \text{if } E_{m+1} \text{ is logically independent on } \{E_1, E_2, \ldots, E_m\}, \\ \in [p', p''] & \text{if } E_{m+1} \text{ is logically dependent on } \{E_1, E_2, \ldots, E_m\}, \end{cases}$$

where p' and p'' are lower and upper probabilities.

The first case (linear dependence) is a special case of logical dependence in which $p' = p''$. Practically all theorems of elementary probability theory belong to the first case. In the second case we say an argument is *probabilistically non-informative*.

As two corollaries we obtain (compare also Fig. 1):

Corollary 1 *(Partial independence from below). If the set of atoms in which the indicator of the conclusion is* 0 *is logically independent, then* $p' = 0$.

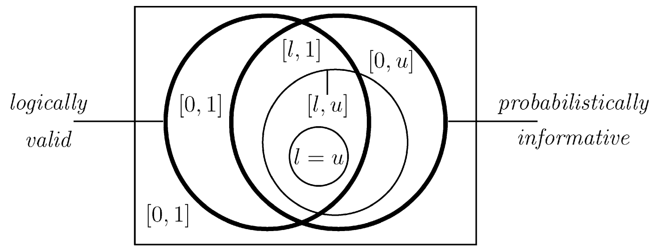

Fig. 1. Classification of probabilistic argument forms. l and u denote whether the lower or upper probability bound of the conclusion, respectively, is constrained by the premise(s). The circle on the left contains argument forms that are logically valid in their non-probabilistic version. The intersection of the bold circles contains the p-valid argument forms. All regions are non-empty, see Table 2 for examples.

Corollary 2 *(Partial independence from above). If the set of atoms in which the indicator of the conclusion is* 1 *is logically independent, then* $p'' = 1$.

One of the best known principles in probability logic is Adams' concept of p-validity [1,2]:

Definition 4 *(Adams' Hauptsatz).* The uncertainty of the conclusion of a valid inference cannot exceed the sum of the uncertainties of its premises.

The uncertainty $u(A)$ is defined by the 1-complement of the corresponding probability, $u(A) = 1 - P(A)$. It may be shown that for unconditional events Adams' Hauptsatz becomes

Corollary 3. *Let* $E_1, E_2, \ldots, E_m$ *be a set of* unconditional *events with probabilities* $p_1, p_2, \ldots, p_m$. *The probability of the conclusion of a valid inference cannot be less than*

$$P(E_{m+1}) = \max\left\{0, 1 - \left(m - \sum_{i=1}^{m} p_i\right)\right\}.$$

The corollary shows that the lower probability is not sensitive (i) to the specific logical form of the premises and (ii) to the order of the probabilities $p_1, \ldots, p_m$. The two properties hold for unconditional events only and reflect the fact that in this case the events are truth functional. Only the lower bounds of the conclusions of those arguments that contain conditional events can be sensitive to the structure of the premises and to the specific pattern of the probability assessment. Algebraically this is associated with the matrix **V** which, in the case of conditionals, does not only contain 0s and 1s, but a pattern of real-valued probabilities.

Material implication is truth functional. If human subjects would interpret IF–THEN as material implication their probability responses in p-valid arguments should be insensitive to the logical form of the premises and to permutations of the probabilities of the premises. There is, however, strong evidence that *human subjects are sensitive to structure and assignment*. We consider this as one of the strongest arguments against the interpretation of the IF–THEN as a material implication.

3. Combining logic and probability in psychology

Recent probabilistic approaches to human deductive reasoning may be classified according to the interpretation of the IF–THEN and according to the relation between the premise(s) and the conclusion. One of the most influential psychological theories of human reasoning is the *mental model* theory [19]. The theory was extended to human probabilistic reasoning [16,20]. The core meaning of the uncertain IF–THEN is postulated to correspond to the probability of the material implication, $P(H \supset E)$. In recent studies on the meaning of the IF–THEN [13,24,27,28], the participants had either to infer the probabilities of the four truth table cases ($E \wedge H$, $E \wedge \neg H$, $\neg E \wedge H$, and $\neg E \wedge \neg H$) from the probability of the IF–THEN (i), or the participants had to infer the probability of the IF–THEN from the probabilities

Table 2
Probability logical argument forms. The logical operators are defined as usual, "$\models$" denotes classical logical truth. The axioms of SYSTEM P are marked by "$*$", derived rules are marked by "†". Derivations of the probability propagation rules of SYSTEM P are in [14], for the other argument forms see [31]. "v" denotes logical validity of the non-probabilistic version of the argument form. "l" and "u" denote whether the lower or the upper probability bound of the conclusion, respectively, is constrained by the probabilities of the premises. "p" denotes whether the argument form is p-valid

Name	Probabilistic version of the argument form	v	l	u	p
LEFT LOGICAL EQUIVALENCE*	$\models (E_1 \equiv E_2), P(E_3 \mid E_1) = x \therefore P(E_3 \mid E_2) = x$	y	y	y	y
PROOF BY CASES†	$P(E_2 \mid E_1) = x, P(E_2 \mid \neg E_1) = y \therefore P(E_2) \in [\min\{x, y\}, \max\{x, y\}]$	y	y	y	y
RIGHT WEAKENING*	$P(E_1 \mid E_3) = x, \models (E_1 \supset E_2) \therefore P(E_2 \mid E_3) \in [x, 1]$	y	y	n	y
MODUS PONENS†	$P(E_2 \mid E_1) = x, P(E_1) = y \therefore P(E_2) \in [xy, 1 - y + xy]$	y	y	y	y
CUT*	$P(E_2 \mid E_1 \wedge E_3) = x, P(E_1 \mid E_3) = y \therefore P(E_2 \mid E_3) \in [xy, 1 - y + xy]$	y	y	y	y
AND†	$P(E_2 \mid E_1) = x, P(E_3 \mid E_1) = y \therefore P(E_2 \wedge E_3 \mid E_1) \in [\max\{0, x + y - 1\}, \min\{x, y\}]$	y	y	y	y
MODUS TOLLENS†	$P(E_2 \mid E_1) = x, P(\neg E_2) = y \therefore P(\neg E_1) \in [\max\{(1 - x - y)/(1 - x), (x + y - 1)/x\}, 1]$	y	y	n	y
CAUTIOUS MONOTONICITY*	$P(E_2 \mid E_1) = x, P(E_3 \mid E_1) = y \therefore P(E_3 \mid E_1 \wedge E_2) \in [\max\{0, (x + y - 1)/x\}, \min\{y/x, 1\}]$	y	y	y	y
OR*	$P(E_3 \mid E_1) = x, P(E_3 \mid E_2) = y \therefore P(E_3 \mid E_1 \vee E_2) \in [xy/(x + y - xy), (x + y - 2xy)/(1 - xy)]$	y	y	y	y
DENYING THE ANTECEDENT	$P(E_2 \mid E_1) = x, P(\neg E_1) = y \therefore P(\neg E_2) \in [(1 - x)(1 - y), 1 - x(1 - y)]$	n	y	y	n
AFFIRMING THE CONSEQUENT	$P(E_2 \mid E_1) = x, P(E_2) = y \therefore P(E_1) \in [0, \min\{y/x, (1 - y)/(1 - x)\}]$	n	n	y	n
TRANSITIVITY	$P(E_2 \mid E_1) = x, P(E_3 \mid E_2) = y \therefore P(E_3 \mid E_1) \in [0, 1]$	y	n	n	n
CONTRAPOSITION	$P(E_2 \mid E_1) = x \therefore P(\neg E_1 \mid \neg E_2) \in [0, 1]; P(\neg E_1 \mid \neg E_2) = x \therefore P(E_2 \mid E_1) \in [0, 1]$	y	n	n	n
MONOTONICITY	$P(E_3 \mid E_1) = x \therefore P(E_3 \mid E_1 \wedge E_2) \in [0, 1]$	y	n	n	n

of the four truth table cases (ii). The interpretation of the $P(\text{IF } H, \text{THEN } E)$ is then easily computed, as the probabilities of all truth table cases are given. The main result was that most subjects interpret the $P(\text{IF } H, \text{THEN } E)$ as a conditional probability, $P(E|H)$, and not as a probability of the material implication, $P(H \supset E)$.

There are at least two ways in which probabilities may enter argument forms. The relation between the premises and the conclusion can be probabilistic (i), or, the inference relation is deductive but some or all premises and the conclusion may be probabilistically valuated (ii). Oaksford, Chater, and Larkin [26] (see also [25]) proposed that the endorsement rate of the conditional inferences is directly proportional to the conditional probability of the conclusion given the categorical premise. The MODUS PONENS, e.g., is evaluated by $P(E|H)$, where "H" denotes the categorical premise, and "E" denotes the conclusion. Liu [23] proposed to conditionalize on both, the categorical premise, and the conditional premise. Thus, in both approaches [23,25,26] the relation between the premises and the conclusion is *probabilistic* and not deductive.

In probability logic a probability is attached to some or all premises and the probability of the conclusion is derived by mathematical methods. Thus, the relation between the premises and the conclusion is *deductive* and not probabilistic. In the coherence approach we consider a set of premises $E_1, \ldots, E_m$ and a conclusion E_{m+1} and assume that there exists a coherent probability assessment $p_1, \ldots, p_m$ for the premises. The probability of the conclusion, p_{m+1}, is derived deductively from the premises. The events (propositions) may be conditional or unconditional. Our own work follows this approach.

Table 2 lists probabilistic versions of well known argument forms. We investigated empirically [29,30,32] a coherence based probabilistic semantics [14] of the basic non-monotonic reasoning SYSTEM P [21]. The rules of SYSTEM P are p-valid [1,2]. A necessary condition for p-validity is that the non-probabilistic versions of the rules are logically valid. Logical validity, however, does not guarantee p-validity. TRANSITIVITY, e.g., is logically valid but not p-valid. We call an argument form "*probabilistically informative*" if the coherent probability interval of its conclusion is not necessarily equal to the unit interval [0, 1]. An inference rule is probabilistically non-informative, if the assignment of the unit interval to its conclusion is necessarily coherent. All p-valid arguments are probabilistically informative (see the classification in Fig. 1).

We empirically investigated rules of SYSTEM P and rules that clearly violate SYSTEM P (see Table 2). We also investigated the probabilistic versions of classical argument forms, like the MODUS PONENS. In the psychology of reasoning classical argument forms were often investigated experimentally. It is interesting to compare results of the probabilistic and the more traditional non-probabilistic argument forms. We give a brief overview of our investigations on SYSTEM P and describe an example study in more detail below.

We translated the non-monotonic inference rules of SYSTEM P into cover-stories. The cover-stories contained the probabilities of the premises. The task of the participants was to infer the probability(-interval) of the conclusions. In all experiments we payed special attention to create an atmosphere of reasoning and to avoid quick guessing. The participants were students of our university. They were tested individually in a quiet room in the department. They were asked to take enough time.

Practically all responses in the LEFT LOGICAL EQUIVALENCE (see Table 2) and in the RIGHT WEAKENING conditions were coherent. For the AND, CAUTIOUS CUT, and the CAUTIOUS MONOTONICITY tasks, more than half of the interval responses were coherent. Moreover, we investigated argument forms that clearly violate SYSTEM P, namely MONOTONICITY, CONTRAPOSITION, and TRANSITIVITY. These argument forms are probabilistically non-informative. Except for TRANSITIVITY, most participants understood that these argument forms are probabilistically non-informative and they inferred wide intervals. We explain the results in the TRANSITIVITY tasks by conversational implicatures. Adams [1] stressed the probabilistic invalidity of the TRANSITIVITY and suggested to interpret TRANSITIVITY in common sense arguments as CUT. If a speaker first utters a premise of the form IF E_1, THEN E_2 and then utters as the second premise IF E_2, THEN E_3, the speaker actually means by the second premise IF E_1 AND E_2, THEN E_3. The speaker does not mention "E_1 AND" to the addressee because E_1 AND is already conversationally implied and "clear" from the context. Thus, we analyzed the data of the TRANSITIVITY tasks as CUT and observed analogue patterns as in the CUT tasks.

Of special interest are the tasks in which all premises are certain. This is the case in those tasks in which the probabilities of the premises are equal to 1. These tasks serve as "control conditions" as they are comparable to the respective non-probabilistic argument forms. In the tasks with certain premises, practically all participants endorse the SYSTEM P rules. The high endorsement rates are comparable to the endorsement rates of the non-probabilistic version of the MODUS PONENS (89–100%; [11]). In the MONOTONICITY task with certain premises the interval responses are large, which means that many participants understand the probabilistic non-informativeness of the MONOTONICITY argument form even in this special condition. In the case of TRANSITIVITY the mean lower bounds are very high. As discussed above, participants might interpret the TRANSITIVITY tasks as CUT tasks.

As an example, we describe a study on the probabilistic versions of the argument form MODUS PONENS (MP), MODUS TOLLENS (MT), DENYING THE ANTECEDENT (DA), and AFFIRMING THE CONSEQUENT (AC). The non-probabilistic MP and MT are logically valid. The non-probabilistic DA and AC are not logically valid. The probabilistic MP and MT are p-valid. The probabilistic DA and AC are not p-valid. We were especially interested to compare the affirmative and the negated versions of these argument forms (i.e., negated conclusions). The four affirmative argument forms (MP, DA, MT, AC) and their negated versions (NMP, NDA, NMT, NAC) are shown in Table 6. The non-probabilistic versions of these argument forms were extensively investigated empirically [11,12]. The non-probabilistic MP is actually endorsed by 89–100%, the MT by 41–81%, the DA by 17–73%, and the non-probabilistic AC is endorsed by 23–75% of the participants [11].

One hundred and twenty students participated in our study on the probabilistic versions. Thirty participants were assigned to each of the four conditions MP and NMP, DA and NDA, MT and NMT, and AC and NAC. Each participant solved three affirmative and three negated arguments. As an example, a MP task had the following form:

> Imagine the following situation. Around Christmas time a certain ski-resort is very busy. This region is very popular among sportsmen, like skiers, snow-boarders, and sledge-riders. Every hour a cable-car brings the sportsmen to the top. About this cable-car we know:
>
> *Exactly* 70% of the *skiers* wear *red caps*.
> *Exactly* 90% of the sportsmen are *skiers*.
> Imagine all the sportsmen in this cable car. How many of these sportsmen wear a *red cap*?

Participants could respond either by a point value or by two interval values. All tasks had a similar structure. Table 3 lists the probabilities presented in the premises, the normative lower and upper bounds, and the participants' mean lower and upper bound responses for the tasks.

In the MP tasks with certain premises (100% in both premises) all thirty participants solved the task correctly and responded "100%". Likewise, all participants solved the negated version of the MODUS PONENS (NMP) correctly and responded "0%". This indicates two things. First, the participants are perfect in the "certain MP" and "certain NMP".

Table 3
Mean lower (LBR) and mean upper bound responses (UBR). P_1 and P_2 denote the probabilities presented in the premises. CLB and CUB denote the normative/coherent bounds. The data of the upper half of the table are taken from [33]

P_1	P_2	CLB	CUB	LBR	UBR	CLB	CUB	LBR	UBR
		MODUS PONENS				NEGATED MODUS PONENS			
1	1	1	1	1	1	.00	.00	.00	.00
.7	.9	.63	.73	.62	.69	.27	.37	.35	.42
.7	.5	.35	.85	.43	.55	.15	.65	.41	.54
		DENYING THE ANTECEDENT				NEG. DENYING THE ANTECEDENT			
1	1	.00	1	.37	.85	.00	1	.01	.53
.7	.2	.20	.44	.19	.42	.56	.80	.52	.76
.7	.5	.15	.65	.25	.59	.35	.85	.33	.65
		MODUS TOLLENS				NEGATED MODUS TOLLENS			
1	1	1	1	.73	.82	.00	.00	.18	.33
.7	.9	.86	1	.46	.72	.00	.14	.20	.41
.7	.5	.29	1	.36	.66	.00	.71	.27	.57
		AFFIRMING THE CONSEQUENT				NEG. AFFIRMING THE CONSEQUENT			
1	1	.00	1	.36	.97	.00	1	.04	.64
.7	.9	.00	.33	.43	.86	.67	1	.10	.48
.7	.5	.00	.71	.34	.77	.29	1	.13	.56

Table 4
Deviance of the number of observed responses falling into the coherent intervals from the expect number assuming a random interval generator. High χ^2 values indicate more than expected coherent responses. If there are too few cases in the coherent interval the χ^2 values are marked by (–)

P_1	.70	.70	.70	.70	.70	.70	.70	.70
P_2	.90	.50	.90	.50	.20	.50	.20	.50
	MP		NMP		DA		NDA	
χ^2	120.33	13.07	85.33	9.60	10.76	.60	6.42	.60
	MT		NMT		AC		NAC	
χ^2	1.72	(–).01	13.88	.98	(–)10.00	(–).92	(–)8.10	(–)11.11

Second, the reliability of our experimental conditions is high. The results agree with the literature. Human subjects are perfectly competent to make MP inferences.

The relation between the number of responses falling into the normatively correct interval and the size of the normative interval is used as a measure of the agreement of the responses and the normative values. We use a simple χ^2 value to express the agreement, $\chi^2 = (f - e)^2/e$, with $f =$ number of participants inferring coherent values, and $e =$ expected number of participants assuming a random response generator. Let the normative interval be $[l, u]$. In step one the random number generator selects a lower response r_l greater than l with probability $1 - l$. In step two it selects a number greater than r_l and less than u with probability $(u - r_l)/(1 - r_l)$. For our purposes it is sufficient to approximate r_l roughly by l. Combining step one and step two by multiplying the two probabilities simplifies to $u - l$ so that $e = N \cdot (u - l)$ where N denotes the number of participants in an experimental condition. Table 4 reports the χ^2 values for the various tasks for the probabilistic premises.

The by far best agreement with the coherent intervals is obtained for the MP and the NMP. There is a significant deviance from normative intervals in the AC and NAC. Data shows that in the first AC task the participants give incoherent responses with too high values and in the NAC with too low values. Matching is one possible explanation, but omitting a negation step in the solution process (see below) is an alternative explanation.

In the DA tasks with "100%" in both premises, fourteen of the thirty participants responded correctly with the unit interval or an interval with a lower boundary very close to zero, $[\leqslant 1, 100]\%$. Practically half of the participants understood that only a non-informative interval can be inferred if each premise is certain.

All participants inferred a probability (interval) of a conclusion $\mathfrak{C}$, $P(\mathfrak{C}) \in [z'_{\mathfrak{C}}, z''_{\mathfrak{C}}]$, and the probability of the associated negated conclusion, $P(\neg\mathfrak{C}) \in [z'_{\neg\mathfrak{C}}, z''_{\neg\mathfrak{C}}]$. To test the conjugacy principle of the interval responses, we

Table 5
Percentages of participants satisfying the conjugacy principle in the MODUS PONENS, DENYING THE ANTECEDENT, MODUS TOLLENS, and the AFFIRMING THE CONSEQUENT conditions (±2% tolerance, $n = 30$ in each condition)

(P_1, P_2)	(1, 1)	(.7, .9)	(.7, .5)	(.7, .2)	(P_1, P_2)	(1, 1)	(.7, .9)	(.7, .5)
MP	100	53	50		MT	67	43	30
DA	67		30	0	AC	77	23	27

checked for each participant whether both $z'_{\mathfrak{C}} + z''_{\neg\mathfrak{C}} = 1$ and $z'_{\neg\mathfrak{C}} + z''_{\mathfrak{C}} = 1$ are satisfied. Table 5 shows the number of participants that exactly satisfy the conjugacy principle in the four tasks and their negated forms. In the MP with 70% and 90% in the premises, for example, 16 of the thirty participants satisfied both conjugacy conditions. Participants with perfect conjugacy show a remarkable sensitivity with respect to the 1-complements in the context of negation.

4. First steps towards a process model of conditional inferences

Evans [12] gives two task features that explain several of the effects observed in classical argument forms, *directionality* and *negativity*. The MODUS PONENS is a *forward* task. The MODUS TOLLENS is a *backward* task. The MP is a forward argument because it requires an inference from the antecedent to the consequent. The MT is a backward argument because it requires an inverse inference, from the consequent to the antecedent. Directionality is best illustrated by a propositional graph. A propositional graph is a directed graph. The verteces represent propositions and the edges between two verteces represent conditionals. We attach probabilities to the edges. The absolute probability of a proposition is represented by an arc without a parent.

Fig. 2 shows a diagram for two affirmative propositions and their negations. The four possible IF–THEN conditionals are represented by the four arcs $A \xrightarrow{y} B$, $A \xrightarrow{1-y} \neg B$,$\neg A \xrightarrow{q} B$, and $\neg A \xrightarrow{1-q} \neg B$. x, y, z, and q denote the probabilities $P(A)$, $P(B|A)$, $P(B)$, and $P(B|\neg A)$, respectively. Dashed arrows are used when the absolute or the conditional probabilities refer to negated propositions. The propositional graph represents the problem space for a class of conditional inference tasks. With such a diagram the premises of the MP are represented by $\xrightarrow{x} A \xrightarrow{y} B$ and the conclusion is represented by $\xrightarrow{z} B$. The inference consists in the removal of the vertex A. Fig. 3 shows the diagrams for the premises of the MP, MT, and the AC.

Non-probabilistic and probabilistic studies have shown that the MT is more difficult than the MP. How can differences like these be explained with the propositional graphs?

We observed that the participants in our experiments were clearly better in lower than upper probability responses. Normatively the lower probability of the conclusion of the MP, z', is the product of the two premise probabilities, $P(A)P(B|A)$ (see Figs. 2 and 3). A process model assumes that human subjects understand that in MP the conclusion is, in any case, less probable than any of its premises and that the lower probability is obtained by taking $100x\%$ of y or $100y\%$ of x. In multiplicative forward chaining, current running results are obtained from iteratively taking a proportion of the last running results. Such an operation is easy to perform intuitively with degrees of belief. Backward processing is sometimes non-informative. In the AC the lower probability is zero. Such results cannot be obtained by "cascaded inference" as in forward inference, see Fig. 3.

We further suppose that negations make inferences difficult so that human subjects prefer to think in terms of affirmative propositions. In an inference graph this requires taking a "detour" and switching from negations to affirmations

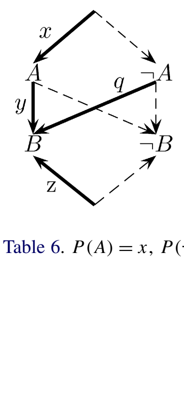

Fig. 2. Probability diagram for the basic argument forms of Table 6. $P(A) = x$, $P(\neg A) = 1 - x$, $P(B|A) = y$, $P(\neg B|A) = 1 - y$, $P(B|\neg A) = q \in [0, 1]$, $P(\neg B|\neg A) = 1 - q$, $P(B) = z$, $P(\neg B) = 1 - z$.

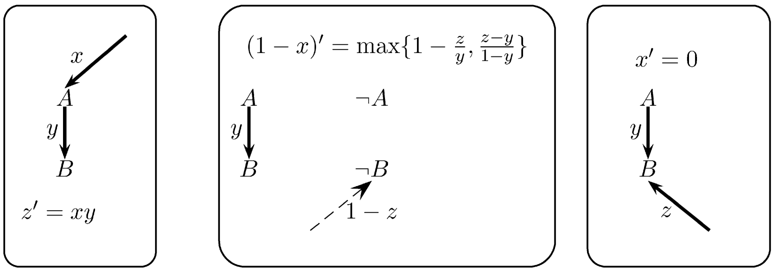

Fig. 3. Left MP, middle MT, right AC. $P(A) = x$, $P(\neg A) = 1 - x$, $P(B|A) = y$, $P(\neg B|A) = 1 - y$, $P(B|\neg A) = q \in [0, 1]$, $P(\neg B|\neg A) = 1 - q$, $P(B) = z$, $P(\neg B) = 1 - z$.

Table 6
MP and negated MP (NMP), DA and negated DA (NDA), MT and negated MT (NMT), AC and negated AC (NAC), the two premises P1 and P2, the conclusion C, $P(B|A) = y$, $P(A) = x$ and $P(B) = z$, $P(\neg A) = 1 - x$, $P(\neg B) = 1 - z$. Compare Fig. 2

	P1	P2	$\therefore$ C	p'	p''
MP	IF A, THEN B	A	B	xy	$1 - x(1 - y)$
NMP	IF A, THEN B	A	$\neg B$	$x(1 - y)$	$1 - xy$
DA	IF A, THEN B	$\neg A$	$\neg B$	$(1 - x)(1 - y)$	$1 - (1 - x)y$
NDA	IF A, THEN B	$\neg A$	B	$(1 - x)y$	$1 - (1 - x)(1 - y)$
MT	IF A, THEN B	$\neg B$	$\neg A$	$\max\{1 - \frac{z}{y}, \frac{z-y}{1-y}\}$	1
NMT	IF A, THEN B	$\neg B$	A	0	$1 - \max\{1 - \frac{z}{y}, \frac{z-y}{1-y}\}$
AC	IF A, THEN B	B	A	0	$\min\{1 - \frac{z}{y}, \frac{z-y}{1-y}\}$
NAC	IF A, THEN B	B	$\neg A$	$1 - \min\{1 - \frac{z}{y}, \frac{z-y}{1-y}\}$	1

and back. It also requires switching 1-complements of probabilities and lower and upper probabilities. This involves more cognitive load. It is easy to lose track on such a solution path or to forget to switch back a 1-complement.

Normatively the upper probability is obtained by the conjugacy principle [38]. The upper probability of an event E is 1 minus the probability of the negation of E, $p''(E) = 1 - p'(\neg E)$. For the upper probability of the MP we need the probability of $\neg B$, the negation of the conclusion. This is obtained by taking from the product $x(1 - y)$ the 1-complement, $z'' = 1 - x(1 - y)$. A process model assumes that these steps are also involved in an analog form in human reasoning. It predicts that the upper probability of the MP is more difficult than the lower one because for its solution more steps are required. The model assumes a strong *preference for affirmative propositions*. In many investigations and in different domains it was observed that negated information requires additional processing efforts, takes more time, leads to more errors, etc. than affirmative information [18].

We distinguish two families of elementary argument forms, the MP and the MT family. Each one has four members. They are obtained by affirming or negating the categorical antecedent or the conclusion. Table 6 gives the lower and upper probabilities for each of the 2×4 argument forms. The conjugacy principle is reflected in the 1-complements of the diagonal entries of the successive argument pairs. The upper probability of the MP, e.g., is equal to 1 minus the lower probability of the NMP. Note the symmetries in the MT family concerning the smaller/greater of two ratios.

The members of the MP family require *forward* processing, those of the MT family *backward* processing. Inferences in the MP family involve *multiplication*, inferences in the MT family *division*. In addition, inferences in the MT family require min/max-decisions. The number and the kind of steps may be used to estimate the difficulty of the argument forms. Obviously the lower probability of MP is especially easy as it requires only one step. MT requires most steps. In the NMP the lower probability of $\neg B$ results from the product of the $x(1 - y)$. The result is obtained by multiplicative chaining (Fig. 3), but this time the 1-complement of the given y value is required. The upper probability for the NMP requires three steps: (i) taking the event-complement of $\neg B$, which is B, and which is given y, (ii) multiplicative chaining, and (iii) taking 1-complement.

By bringing probability, logic and psychology together we have tried to improve the understanding of human reasoning. We have approached the area on several routes simultaneously, including the selection of an appropriate normative framework, running experiments, and modeling of cognitive representations and reasoning processes. The investigation of human reasoning is a highly interdisciplinary endeavor.

References

[1] E.W. Adams, The Logic of Conditionals, Reidel, Dordrecht, 1975.
[2] E.W. Adams, A Primer of Probability Logic, CSLI, Stanford, 1998.
[3] M. Baioletti, A. Capotorti, L. Galli, S. Tognoloni, F. Rossi, B. Vantaggi, CkC (Check Coherence package; version e4, June, 2007). Download: http://www.dipmat.unipg.it/~upkd/paid/software.html.
[4] B. Bouchon-Meunier, G. Coletti, C. Marsala, Conditional possibility and necessity, in: B. Bouchon-Meunier, J. Gutiérrez-Ríos, L. Magdalena, R.R. Yager (Eds.), Technologies for Constructing Intelligent Systems: Tools, Springer, Berlin, 2002, pp. 59–72.
[5] R.M.J. Byrne, The Rational Imagination. How People Create Alternatives to Reality, Bradford, MIT Press, Cambridge, MA, 2005.
[6] G. Coletti, R. Scozzafava, Probabilistic Logic in a Coherent Setting, Kluwer, Dordrecht, 2002.
[7] G. Coletti, R. Scozzafava, Conditional probability, fuzzy sets, and possibility: A unifying view, Fuzzy Sets and Systems 144 (2004) 227–249.
[8] G. Coletti, R. Scozzafava, Conditioning in a coherent setting: Theory and applications, Fuzzy Sets and Systems 155 (2005) 26–49.
[9] B. De Finetti, Theory of Probability, vols. 1, 2, John Wiley & Sons, Chichester, 1974. Original work published 1970.
[10] G.J.W. Dorn, Popper's law of the excess of the probability of the conditional over the conditional probability, Conceptus 26 (67) (1992) 6–61.
[11] J.St.B.T. Evans, S.E. Newstead, R.M.J. Byrne, Human Reasoning, Erlbaum, Hove, 1993.
[12] J.St.B.T. Evans, The Psychology of Deductive Reasoning, Routledge, London, 1982.
[13] J.St.B.T. Evans, S.H. Handley, D.E. Over, Conditionals and conditional probability, Journal of Experimental Psychology: Learning, Memory, and Cognition 29 (2003) 321–355.
[14] A. Gilio, Probabilistic reasoning under coherence in System P, Annals of Mathematics and Artificial Intelligence 34 (2002) 5–34.
[15] A. Gilio, S. Ingrassia, Totally coherent set-valued probability assessments, Kybernetika 34 (1) (1998) 3–15.
[16] V. Girotto, P.N. Johnson-Laird, The probability of conditionals, Psychologia 47 (2004) 207–225.
[17] I.R. Goodman, H.T. Nguyen, E.A. Walker, Conditional Inference and Logic for Intelligent Systems. A Theory of Measure-Free Conditioning, North-Holland, Amsterdam, 1991.
[18] L.R. Horn, A Natural History of Negation, CSLI Publications, Stanford, 2001.
[19] P.N. Johnson-Laird, R.M.J. Byrne, Deduction, Erlbaum, Hillsdale, 1991.
[20] P.N. Johnson-Laird, V. Girotto, P. Legrenzi, M.S. Legrenzi, Naive probability: A mental model theory of extensional reasoning, Psychological Review 106 (1) (1999) 62–88.
[21] S. Kraus, D. Lehmann, M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, Artificial Intelligence 44 (1990) 167–207.
[22] F. Lad, Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction, Wiley, New York, 1996.
[23] I.-M. Liu, Conditional reasoning and conditionalization, Journal of Experimental Psychology: Learning, Memory, and Cognition 29 (4) (2003) 694–709.
[24] I.-M. Liu, K.-C. Lo, J.-T. Wu, A probabilistic interpretation of 'If–Then', The Quarterly Journal of Experimental Psychology 49 (A) (1996) 828–844.
[25] M. Oaksford, N. Chater, Bayesian Rationality: The Probabilistic Approach to Human Reasoning, Oxford University Press, Oxford, 2007.
[26] M. Oaksford, N. Chater, J. Larkin, Probabilities and polarity biases in conditional inference, Journal of Experimental Psychology: Learning, Memory, and Cognition 26 (2000) 883–899.
[27] K. Oberauer, O. Wilhelm, The meaning(s) of conditionals: Conditional probabilities, mental models and personal utilities, Journal of Experimental Psychology: Learning, Memory, and Cognition 29 (2003) 680–693.
[28] D.E. Over, C. Hadjichristidis, J.S. Evans, S.J. Handley, S.A. Sloman, The probability of causal conditionals, Cognitive Psychology 54 (2007) 62–97.
[29] N. Pfeifer, G.D. Kleiter, Nonmonotonicity and human probabilistic reasoning, in: Proceedings of the 6th Workshop on Uncertainty Processing, Hejnice, September 24–27, 2003, pp. 221–234.
[30] N. Pfeifer, G.D. Kleiter, Coherence and nonmonotonicity in human reasoning, Synthese 146 (1–2) (2005) 93–109.
[31] N. Pfeifer, G.D. Kleiter, Inference in conditional probability logic, Kybernetika 42 (2006) 391–404.
[32] N. Pfeifer, G.D. Kleiter, Is human reasoning about nonmonotonic conditionals probabilistically coherent? in: Proceedings of the 7th Workshop on Uncertainty Processing, Mikulov, September 16–20, 2006, pp. 138–150.
[33] N. Pfeifer, G.D. Kleiter, Human reasoning with imprecise probabilities: Modus ponens and Denying the antecedent, in: 5th International Symposium on Imprecise Probability: Theories and Applications, Prague, Czech Republic, 16–19 July 2007, pp. 347–356.
[34] K.R. Popper, A set of independent axioms for probability, Mind 47 (1938) 275–277.
[35] F.P. Ramsey, Truth and probability (1926), in: D.H. Mellor (Ed.), Philosophical Papers by F.P. Ramsey, Cambridge University Press, Cambridge, 1994, pp. 58–100.
[36] F.P. Ramsey, General propositions and causality (1929), in: D.H. Mellor (Ed.), Philosophical Papers by F.P. Ramsey, Cambridge University Press, Cambridge, 1994, pp. 145–163.
[37] A. Rényi, On a new axiomatic theory of probability, Acta Mathematica Academiae Scientiarum Hungaricae 6 (1955) 285–335.
[38] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991.

Chapter 4

General conclusion

In this thesis I provided new clarifications and solutions to a number of central epistemological questions. I proposed coherence based probability logic as a fruitful and unified rationality framework for various epistemological problems of uncertain inference. Specifically, I proposed a new measure of argument strength that accounts for the logical structure of the premise set and that is sensitive to the imprecision and probability of the conclusion (Section 3.1, p. 25). As mentioned in Section 2.1, future empirical research is needed to investigate the psychological plausibility of the proposed measure.

The advantage of basing the new measure on a deductive consequence relation could be exploited in current debates in the philosophy of science. Specifically, the proposed measure in Section 3.1 can be used as a cornerstone of a new theory of confirmation. The key idea is that data (i. e., the premises $\mathcal{P}$) plus background knowledge $\mathcal{K}$ confirm a hypothesis (i. e., the conclusion $\mathcal{C}$), if the $\mathcal{P}$ probabilities plus $\mathcal{K}$ raise $P(\mathcal{C})$ and simultaneously decrease the imprecision of $P(\mathcal{C})$.

In the context of argumentation theory, coherence based probability logic has been shown to be fruitful not only in elaborating a measure of argument strength, but also in the analysis of fallacies. In Pfeifer (2008) I investigate selected reasoning fallacies. Specifically, I argue that the *argumentum ad ignorantiam*, *affirming the consequent*, *denying the antecedent*[1] and the *conjunction fallacy* are not fallacies *per se*. Rather, the violation of coherence is a sufficient condition for an argument to be fallacious.

In Section 3.2 (p. 41) I analyze formally the semantics of negated apparently self-contradictory conditionals in the context of Aristotle's theses and propose the conditional event interpretation of conditionals as a fruitful and viable alternative to connexive logics (McCall, 1966; Angell, 2002). Moreover, two experiments on both versions of Aristotle's Thesis demonstrate the psychological plausibility of the proposed semantics. The conditional event interpretation fits with common sense and neither the wide nor the narrow scope reading of the negation of material conditionals are supported by the empirical data. This provides new and empirical justifications of the criticism of using material conditionals for formalizing common sense arguments. While the paradoxes of the material conditional are popular in the discussions about the semantics of conditionals, Aristotle's theses are often ignored. Section 3.2 illustrates the theoretical and empirical fruitfulness of Aristotle's thesis for the first time.

A probabilistic interpretation of two paradoxes of the material conditional is presented in Section 3.3 (p. 59). While advocates of the material conditional often explain the paradoxical nature of the material conditional away by using pragmatic arguments (e.g., in the sense of Grice (1975)), I show that pragmatics is not necessary to avoid these paradoxes. The conditional event interpretation of conditionals is free of these paradoxes of the material conditional and does not employ any pragmatic considerations. Thus, my analysis is much more parsimonious as it is elaborated in purely semantical terms. Two reported original experiments substantiate the formal analysis and highlight the empirical plausibility. Future work

[1]See Section 2.5, p. 23 above.

is needed to study the case of the paradoxes in the context of additional probabilistic or logical knowledge.

I am not arguing that all indicative paradoxes are formalizable in purely probabilistic terms. The irrelevant inference from A to $A \vee B$ is interpreted in probability logic as

$$\begin{array}{ll} (1) & P(A) = x \\ \hline (2) & P(A \vee B) \in [x, 1] \end{array} \quad .$$

If x is a point value and $x < 1$, then the imprecision of the conclusion (2) is trivially higher than the premise imprecision (1). This may be interpreted as a violation of the Gricean maxim of Quantity (Grice, 1975). Moreover, experimental data taken form a dice task suggest that the majority of people's degree of belief is

$$P(\text{An even number shows up} \mid \text{Number 2 shows up}) = 1$$

after throwing randomly a fair die. This is a coherent and informative assessment. However, if the conditioned event (i. e., the consequent of the conditional) is replaced by "Number 2 shows up $\vee$ Number 4 shows up", then the observed modal degree of belief is equal to zero. The zero responses may be interpreted as an expression that the inference is irrelevant (Fugard, Pfeifer, & Mayerhofer, 2011). (There was no such possibility in the response format for expressing this kind of irrelevance.) Further research is needed to clarify this point formally and empirically. Formal clarifications may be obtained by adding pragmatic considerations or by imposing meta-logical conclusion relevance conditions in the sense of Schurz (1991). Empirical clarifications may be obtained by using improved response formats and by conducting interviews with the participants immediately after they have solved the tasks.

Section 3.4 (p. 76) investigates how defeasible and nonmonotonic inferences are justified within a probabilistic framework. A series of psychological experiments shows that the cautious and nonmonotonic rules of System P are corroborated by the participants. However, just those argument forms that imply monotonicity—which are the monotonic counterparts to selected System P rules—are those that are not corroborated by the participants. This speaks for the psychological plausibility of the basic rationality postulates of System P. Formally, coherence based probability logic adequately formalizes the paradigm example of nonmonotonic reasoning in Section 3.4 (i. e., the *Tweety case*, see also the updated discussion in Section 2.4, page 19 above). While System P has been investigated empirically, future work is needed to investigate weaker systems than P (e.g., System O by Hawthorne and Makinson (2007)) and stronger systems than P (e.g., System R by Lehmann and Magidor (1992) or System Z by Goldszmidt and Pearl (1996)).

Section 3.5 (p. 76) studies general properties of uncertain argument forms and the interrelations among logical validity, Adams' p-validity (1975, 1998) and probabilistic informativeness. Moreover, I surveyed a set of central inference rules of probability logic in the light of these general properties. I discussed the advantage of using the tightest coherent lower and upper probability bounds and probabilistic informativeness over the use of p-validity in formal epistemology and the psychology of reasoning. The line of research in Section 3.5 should be continued by studying the impact of relevance and irrelevance assumptions in the premises. This is especially interesting in the context of imprecise probabilities: While independence and relevance are symmetric relations in the special case of precise probabilities, (ir)relevance is not symmetric in the general case of imprecise probabilities (Cozman & Walley, 2005).

All sections of my thesis advance an empirically informed naturalized formal epistemology of uncertain reasoning. Moreover, sections 3.1–3.5 investigate paradigm examples for extending the current domain of experimental philosophy. Contrary to some of the shallow, often inventory-based empirical methodologies of current experimental philosophical work, I make a strong case for properly designed and carefully carried out experiments. Moreover, I advanced the practice of instructing the participants to evaluate the task comprehensibility and the understandability of the task material after the experiments. High ranks in these evaluations speak for the carefully formulated instructions and task materials in the reported experiments.

Further directions of future work include formal and empirical investigations on counterfactual conditionals as well as on causal reasoning and on abductive reasoning. Moreover, the investigated carrier structure of probabilities is composed of propositional logical formulæ, enriched with the conditional event. Thus, extending the proposed coherence based probability logic towards a formalism dealing with quantifiers is a further direction of future research. Here, Angelo Gilio, Giuseppe Sanfilippo and I already obtained interesting results in developing a coherence based probability semantics for Aristotelian syllogisms. Specifically, we probabilized the frequency semantics I presented in Pfeifer (2006a) and obtained various probabilistic notions of quantifiers, existential imports, and generalizations of the traditional square of oppositions (Pfeifer, Sanfilippo, & Gilio, 2010; Pfeifer, 2011a).

A further frontier of future research involves investigations on the relations among the coherence approach and other uncertainty measures like possibility theory, Dempster-Shafer belief functions or fuzzy sets. In a recently granted research project (*German Research Foundation*; co-applicant: Gabriele Kern-Isberner) we will study the relations between coherence based probability logic, conditional structures, and maximum entropy approaches aiming to establish new rationality norms for conditionals from a formal and an empirical point of view.

References

Adams, E. W. (1975). *The logic of conditionals*. Dordrecht: Reidel.

Adams, E. W. (1998). *A primer of probability logic*. Stanford: CSLI.

Angell, R. B. (2002). *A-logic*. Lanham: University Press of America.

Benferhat, S., Bonnefon, J.-F., & Da Silva Neves, R. (2005). An overview of possibilistic handling of default reasoning, with experimental studies. *Synthese*, *1-2*, 53-70.

Biazzo, V., & Gilio, A. (2000). A generalization of the fundamental theorem of De Finetti for imprecise conditional probability assessments. *International Journal of Approximate Reasoning*, *24*(2-3), 251-272.

Bonnefon, J. F., & Politzer, G. (2010). Paradoxes, mental models and one paradox of the material conditional. *Mind & Language*, *26*(2), 141–155.

Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Oxford University Press.

Brewka, G. (1991). *Nonmonotonic reasoning: Logical foundations of commonsense*. Cambridge: Cambridge University Press.

Coletti, G., & Scozzafava, R. (2002). *Probabilistic logic in a coherent setting*. Dordrecht: Kluwer.

Cozman, F. G., & Walley, P. (2005). Graphoid properties of epistemic irrelevance and independence. *Annals of Mathematics and Artificial Intelligence*, *45*(1-2), 173-195.

Crupi, V., Tentori, K., & Gonzales, M. (2007). On Bayesian measures of confirmation. *Philosophy of Science*, *74*, 229-252.

De Finetti, B. (1974). *Theory of probability* (Vols. 1, 2). Chichester: John Wiley & Sons. (Original work published 1970)

De Finetti, B. (1980). Foresight: Its logical laws, its subjective sources (1937). In H. J. Kyburg & H. E. Smokler (Eds.), *Studies in subjective probability* (p. 55-118). Huntington, New York: Robert E. Krieger Publishing Company.

Douven, I. (2008). The evidential support theory of conditionals. *Synthese*, *164*, 19–44.

Douven, I. (2012). Learning conditional information. *Mind & Language*, *27*(3), 239–263.

Douven, I., & Meijs, W. (2007). Measuring coherence. *Synthese*, *156*(3), 405-425.

Elio, R., & Pelletier, F. J. (1993). Human benchmarks on ai's benchmark problems. In *Proceedings of the 15th annual conference of the cognitive science society* (p. 406-411). Boulder, CO: Morgan Kaufmann.

Evans, J. St. B. T. (in press). Questions and challenges to the new psychology of reasoning. *Thinking & Reasoning*.

Feldman, R. (2008). Naturalized epistemology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2008 ed.). `http://plato.stanford.edu/archives/fall2008/entries/epistemology-naturalized/`.

Feltz, A. (2009). Experimental philosphy. *Analyse & Kritik*, *2*, 201–219.

Fischhoff, B., & Kadvany, J. (2011). *Risk. A very short introduction*. Oxford: Oxford University Press.

Ford, M. (2004). System LS: A three-tiered nonmonotonic reasoning system. *Computational Intelligence*, *20*(1), 89-108.

Ford, M., & Billington, D. (2000). Strategies in human nonmonotonic reasoning. *Computational Intelligence*, *16*(3), 446-468.

Fugard, A. J. B., Pfeifer, N., & Mayerhofer, B. (2011). Probabilistic theories of reasoning need pragmatics too: Modulating relevance in uncertain conditionals. *Journal of Pragmatics*, *43*, 2034-2042.

Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2009). How people interpret an uncertain If. In T. Kroupa & J. Vejnarova (Eds.), *Proceedings of the 8th workshop on uncertainty processing* (p. 80-91). Liblice, CZ.

Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts towards the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 635–648.

Galavotti, M. C. (Ed.). (2008). *Bruno de Finetti. Radical probabilist*. London: College Publications.

Gilio, A. (2002). Probabilistic reasoning under coherence in System P. *Annals of Mathematics and Artificial Intelligence*, *34*, 5-34.

Goldszmidt, M., & Pearl, J. (1996). Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence*, *84*, 57-112.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3: Speech acts.). New York: Seminar Press.

Haenni, R., Romeijn, J.-W., Wheeler, G., & Williamson, J. (2011). *Probabilistic logics and probabilistic networks*. Dordrecht: Springer.

Hahn, U., & Oaksford, M. (2006). A normative theory of argument strength. *Informal Logic*, *26*, 1-22.

Hailperin, T. (2000). Probability semantics for quantifier logic. *Journal of Philosophical Logic*, *29*, 207-239.

Hailperin, T. (2011). *Logic with a probability semantics. including a probability semantics to some philosophical problems*. Bethlehem: Lehigh University Press.

Hájek, A. (2011). Interpretations of probability. In E. N. Zalta

(Ed.), *The Stanford encyclopedia of philosophy* (December 2011 ed.). `http://plato.stanford.edu/entries/probability-interpret/`. (retrieved February 2012)

Hawthorne, J., & Makinson, D. (2007). The quantitative/qualitative watershed for rules of uncertain inference. *Studia Logica*, *86*, 247-297.

Hendricks, V. F. (2006). *Mainstream and formal epistemology*. Cambridge: Cambridge University Press.

Huber, F., & Schmidt-Petri, C. (Eds.). (2009). *Degrees of belief*. Dordrecht: Springer.

Knobe, J., & Nichols, S. (Eds.). (2008). *Experimental philosophy*. Oxford: Oxford University Press.

Kraus, S., Lehmann, D., & Magidor, M. (1990). Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, *44*, 167-207.

Lad, F. (1996). *Operational subjective statistical methods: A mathematical, philosophical, and historical introduction*. New York: Wiley.

Lehmann, D., & Magidor, M. (1992). What does a conditional knowledgebase entail? *Artificial Intelligence*, *55*(1), 1-60.

McCall, S. (1966). Connexive implication. *Journal of Symbolic Logic*, *31*, 415-433.

McCarthy, J. (1977). Epistemological problems of artificial intelligence. In *Ijcai'77: Proceedings of the 5th international joint conference on artificial intelligence* (pp. 1038–1044). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Oaksford, M., & Chater, N. (2009). Précis of "Bayesian rationality: The probabilistic approach to human reasoning". *Behavioral and Brain Sciences*, *32*, 69-120.

Over, D. E., Hadjichristidis, C., Evans, J. St. B. T., Handley, S. J., & Sloman, S. (2007). The probability of causal conditionals. *Cognitive Psychology*, *54*, 62-97.

Pelletier, F. J., & Elio, R. (1997). What should default reasoning be, by default? *Computational Intelligence*, *13*, 165-187.

Pfeifer, N. (2002). *Psychological investigations on human nonmonotonic reasoning with a focus on System P and the conjunction fallacy*. Unpublished master's thesis, Institut für Psychologie, Universität Salzburg. (`http://www.pfeifer-research.de/pdf/PsyHNmrNP.pdf`)

Pfeifer, N. (2006a). Contemporary syllogistics: Comparative and quantitative syllogisms. In G. Kreuzbauer & G. J. W. Dorn (Eds.), *Argumentation in Theorie und Praxis: Philosophie und Didaktik des Argumentierens* (p. 57-71). Wien: LIT.

Pfeifer, N. (2006b). *On mental probability logic*. Unpublished doctoral dissertation, Department of Psychology, University of Salzburg. (The abstract is published in *The Knowledge Engineering Review*, 2008, 23, pp. 217-226; `http://www.pfeifer-research.de/pdf/diss.pdf`)

Pfeifer, N. (2007). Rational argumentation under uncertainty. In G. Kreuzbauer, N. Gratzl, & E. Hiebl (Eds.), *Persuasion und Wissenschaft: Aktuelle Fragestellungen von Rhetorik und Argumentationstheorie* (p. 181-191). Wien: LIT.

Pfeifer, N. (2008). A probability logical interpretation of fallacies. In G. Kreuzbauer, N. Gratzl, & E. Hiebl (Eds.), *Rhetorische Wissenschaft: Rede und Argumentation in Theorie und Praxis* (pp. 225–244). Wien: LIT.

Pfeifer, N. (2011a, September, 10–11). *Coherence based probability logic: Philosophical and psychological applications.* Talk given at "The Fifth Workshop on Combining Probability and Logic (progic2011)" workshop, New York (U.S.A.).

Pfeifer, N. (2011b). Systematic rationality norms provide research roadmaps and clarity. Commentary on Elqayam & Evans: Subtracting "ought" from "is": Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences*, *34*, 263–264.

Pfeifer, N. (2012). Experiments on Aristotle's Thesis: Towards an experimental philosophy of conditionals. *The Monist*, *95*(2), 223–240.

Pfeifer, N., & Douven, I. (submitted). Toward an experimental philosophy of uncertain reasoning. *Synthese*.

Pfeifer, N., & Kleiter, G. D. (2002). Experiments on nonmonotonic reasoning. the coherence of human probability judgments. In H. Leitgeb & G. Schurz (Eds.), *Proceedings of the* 1^{st} *salzburg workshop on paradigms of cognition.* Universität Salzburg.

Pfeifer, N., & Kleiter, G. D. (2003). Nonmonotonicity and human probabilistic reasoning. In *Proceedings of the* 6^{th} *workshop on uncertainty processing* (p. 221-234). Hejnice: September 24–27^{th}, 2003.

Pfeifer, N., & Kleiter, G. D. (2005). Coherence and nonmonotonicity in human reasoning. *Synthese*, *146*(1-2), 93-109.

Pfeifer, N., & Kleiter, G. D. (2006a). Inference in conditional probability logic. *Kybernetika*, *42*, 391-404.

Pfeifer, N., & Kleiter, G. D. (2006b). Is human reasoning about nonmonotonic conditionals probabilistically coherent? In *Proceedings of the* 7^{th} *workshop on uncertainty processing* (p. 138-150). Mikulov: September 16–20^{th}, 2006.

Pfeifer, N., & Kleiter, G. D. (2007). Human reasoning with imprecise probabilities: Modus ponens and Denying the antecedent. In G. De Cooman, J. Vejnarová, & M. Zaffalon (Eds.), 5^{th} *International Symposium on Imprecise Probability: Theories and Applications* (p. 347-356). Prague, Czech Republic: SIPTA.

Pfeifer, N., & Kleiter, G. D. (2009). Framing human inference by coherence based probability logic. *Journal of Applied Logic*, *7*(2), 206–217.

Pfeifer, N., & Kleiter, G. D. (2011). Uncertain deductive reasoning. In K. Manktelow, D. E. Over, & S. Elqayam (Eds.), *The science of reason:*

A Festschrift for Jonathan St. B.T. Evans (p. 145-166). Hove: Psychology Press.

Pfeifer, N., Sanfilippo, G., & Gilio, A. (2010, December, 8–9). *Coherent probabilistic quantification, existential import and aristotelian syllogistics.* Talk given at the "Vague Quantities and Vague Quantifiers (VQ2)" workshop, Berlin (Germany).

Phillips, J. (2011). *The experimental philosophy webpage.* `http://experimentalphilosophy.org/ExperimentalPhilosophy.html`. (retrieved February 2012)

Quine, W. V. O. (1969). Epistemology naturalized. In W. V. O. Quine (Ed.), *Ontological relativity and other essays* (pp. 69–90). New York: Columbia University Press.

Rips, L. J. (2001). Two kinds of reasoning. *Psychological Science*, *12*(2), 129-134.

Schurz, G. (1991). Relevant deduction: From solving paradoxes towards a general theory. *Erkenntnis*, *35*, 391–437.

Schurz, G. (2005). Non-monotonic reasoning from an evolution-theoretic perspective: Ontic, logical and cognitive foundations. *Synthese*, *1-2*, 37-51.

Schurz, G. (2011). Tweety, or why probabilism and even Bayesianism need objective and evidential probabilities. In D. Dieks, W. J. González, S. Hartmann, M. Stöltzner, & M. Weber (Eds.), *Probabilities, laws, and structures* (Vol. 3, pp. 57–74). Dordrecht: Springer.

Spohn, W. (Ed.). (2009). *Causation, coherence and concepts*. Dordrecht: Springer.

Vogel, C. (1996). Human reasoning with negative defaults. In D. Gabbay & H. J. Ohlbach (Eds.), *Practical reasoning, lecture notes in artificial intelligence, 1085* (p. 606-621). Berlin: Springer.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. London:

Chapman and Hall.

Appendix A

Characterization theorem of coherence

Appendix A reproduces the characterization theorem of coherence (Coletti & Scozzafava, 2002, Theorem 4, p. 81). For an informal characterization of coherence *see Chapter 1 on p. 10 above.*

Let $\mathscr{C}$ be an *arbitrary* family of conditional events and consider, for every $n \in \mathbb{N}$, a *finite* subfamily

$$\mathscr{F} = \{E_1|H_1, \ldots, E_n|H_n\} \subseteq \mathscr{C}\,;$$

we denote by $\mathscr{A}$ the set of *atoms* A_r generated by the (unconditional) events $E_1, H_1, \ldots, E_n, H_n$ and by $\mathscr{G}$ the *algebra* spanned by them. For an *assessment* on $\mathscr{C}$ given by a real function P, the following three statements are equivalent:

1. P is a *coherent* conditional probability on $\mathscr{C}$;

2. for every $n \in \mathbb{N}$ and for every *finite* subset $\mathscr{F} \subseteq \mathscr{C}$ there exists a sequence of *compatible* systems, with unknowns $x_r^\alpha \geq 0$,

$$(S_\alpha)\begin{cases}\displaystyle\sum_{A_r\subseteq E_i\wedge H_i} x_r^\alpha = P(E_i|H_i)\sum_{A_r\subseteq H_i} x_r^\alpha\,, \\ \left[\text{if } \sum_{A_r\subseteq H_i} x_r^{\alpha-1}=0,\ \alpha\geq 1\right] \qquad (i=1,2,\ldots,n) \\ \displaystyle\sum_{A_r\subseteq H_0^\alpha} x_r^\alpha = 1\end{cases}$$

with $\alpha = 0,1,2,\ldots,k \leq n$, where $H_0^0 = H_0 = H_1 \vee \ldots \vee H_n$ and H_0^α denotes, for $\alpha \geq 1$, the union of the H_i's such that $\sum_{A_r\subseteq H_i} x_r^{\alpha-1} = 0$;

3. for every $n \in \mathbb{N}$ and for every *finite* subset $\mathscr{F} \subseteq \mathscr{C}$ there exists (at least) a *class* of (coherent) probabilities $\{P_0^{\mathscr{F}}, P_1^{\mathscr{F}}, \ldots, P_h^{\mathscr{F}}\}$, each probability $P_\alpha^{\mathscr{F}}$ being defined on a suitable subset $\mathscr{A}_\alpha \subseteq \mathscr{A}_0$ (with $\mathscr{A}_{\alpha'} \subseteq \mathscr{A}_{\alpha''}$ for $\alpha' > \alpha''$ and $P_{\alpha''}^{\mathscr{F}}(A_r) = 0$ if $A_r \in \mathscr{A}_{\alpha'}$) such that for every $G \in \mathscr{G}, G \neq \emptyset$, there is a unique $P_\alpha^{\mathscr{F}}$, with

$$\sum_{\substack{r\\A_r\subseteq G}} P_\alpha^{\mathscr{F}}(A_r) > 0\,; \tag{A.1}$$

moreover, for every $E_i|H_i \in \mathscr{F}$ there exists a unique $P_\beta^{\mathscr{F}}$ satisfying (A.1) with $G = H_i$ and $\alpha = \beta$, and $P(E_i|H_i)$ is *represented* in the form

$$P(E_i|H_i) = \frac{\sum_{A_r\subseteq E_i\wedge H_i} P_\beta^{\mathscr{F}}(A_r)}{\sum_{A_r\subseteq H_i} P_\beta^{\mathscr{F}}(A_r)}\,. \tag{A.2}$$

Appendix B

Proofs of the paradoxes

B.1 Paradox 1

The proof of the probabilistic non-informativeness of PARADOX 1,

From $P(B) = 1$ *infer* $P(B|A) \in [0, 1]$

consists of two parts. In part I we show that $P(B|A)$ may be equal to zero if $P(B) = 1$, and in part II we show that $P(B|A)$ may be equal to one if $P(B) = 1$. Ω denotes the *certain event*[1] and $\forall A(P(A) = P(A|\Omega))$.

Part I: Proof that $P(B|A)$ may be equal to zero if $P(B) = 1$

the following assessment on the list of conditional events $\mathscr{C} = \{B|\Omega, B|A\}$ coherent?

$$(0.1) \quad P(B|\Omega) = 1$$

$$(0.2) \quad P(B|A) = 0$$

[1]The certain event is equivalent to the disjunction of all n atoms, $\Omega \Leftrightarrow A_1 \vee \cdots \vee A_n$.

The set of the atoms $\mathscr{A}_0 = \{A_1, \ldots, A_4\}$ is generated by the following list of atoms A_i:

$$\begin{array}{lccccc} A_1 & \Leftrightarrow & A & \wedge & B \\ A_2 & \Leftrightarrow & A & \wedge & \neg B \\ A_3 & \Leftrightarrow & \neg A & \wedge & B \\ A_4 & \Leftrightarrow & \neg A & \wedge & \neg B \end{array}$$

x_i^α denotes the (unknown) probability value of the atom A_i. The (unconditional) probability function that assigns x_i^α to A_i is denoted by $P_\alpha(A_i) = x_i^\alpha$. The index α indicates that the probability function and the probability value are always relative to the respective system (S_α) in the sequence of the systems. The first system (S_0) is the following:

$$(S_0)\begin{cases} (1) & (x_1^0 + x_3^0) = P(B|\Omega)(x_1^0 + x_2^0 + x_3^0 + x_4^0) \\ (2) & x_1^0 = P(B|A)(x_1^0 + x_2^0) \\ (3) & x_1^0 + x_2^0 + x_3^0 + x_4^0 = 1 \\ (4) & \forall i(x_i^0 \geq 0) \end{cases}$$

In the next steps we try to transform the information given in (S_0) such that the probability values of the atoms x_i^0 are equal to zero.

$$\begin{array}{lll} (5) & x_2^0 + x_4^0 = 0 & (1), (0.1) \\ (6) & x_1^0 = 0 & (2), (0.2) \\ (7) & x_3^0 = 1 & (3), (4-6) \end{array}$$

$P_0(B|\Omega) = 1$ is satisfied, since $P_0(B|\Omega) = \frac{x_1^0 + x_3^0}{x_1^0 + x_2^0 + x_3^0 + x_4^0} = \frac{0+1}{0+0+1+0} = 1$. Since x_3^0 is not necessarily equal to zero, we can delete equation (1) and construct the next system:[2]

[2]The condition "*if* $\sum_{\substack{r \\ A_r \subseteq H_i}} x_r^{\alpha-1} = 0,\ \alpha \geq 1$" is not satisfied because step (7) states that $x_3^0 = 1$ (see the second statement of the characterization theorem).

$$(S_1)\begin{cases}(1') & x_1^1 = P(B|A)(x_1^1 + x_2^1)\\ (2') & x_1^1 + x_2^1 = 1\\ (3') & \forall i(x_i^1 \geq 0)\end{cases}$$

$$(4') \quad x_1^1 = 0 \quad (1'), (0.1)$$

$$(5') \quad x_2^1 = 1 \quad (2'), (3', 4')$$

$P_1(B|A) = 0$ is satisfied, since $P_1(B|A) = \frac{x_1^1}{x_1^1+x_2^1} = \frac{0}{0+1} = 0$. Therefore, the assessment (0.1) and (0.2) on $\mathscr{C}$ is coherent. The probability assessment (0.1) and (0.2), however, violates the classical (Kolmogorov) probability axioms. The classical approach defines conditional probability by the fraction of unconditional probabilities. If $P(B) = 1$, then $P(B|A) = \frac{P(A \wedge \neg B)}{P(A)} = \frac{P(A)}{P(A)}$. $P(B|A) = 1$ if $P(A) > 0$, otherwise $P(B|A)$ is undefined. This violates (0.2).

Part II: Proof that $P(B|A)$ may be equal to one if $P(B) = 1$

Is the following assessment on the list of conditional events $\mathscr{C} = \{B|\Omega, B|A\}$ coherent?

$$(0.1) \quad P(B|\Omega) = 1$$

$$(0.2) \quad P(B|A) = 1$$

The set of the atoms $\mathscr{A}_0 = \{A_1, \ldots, A_4\}$ is generated by the following list of atoms A_i:

$$\begin{array}{lcccc} A_1 & \Leftrightarrow & A & \wedge & B\\ A_2 & \Leftrightarrow & A & \wedge & \neg B\\ A_3 & \Leftrightarrow & \neg A & \wedge & B\\ A_4 & \Leftrightarrow & \neg A & \wedge & \neg B\end{array}$$

The system (S_0) is:

$$(S_0)\begin{cases}(1) & (x_1^0 + x_3^0) = P(B|\Omega)(x_1^0 + x_2^0 + x_3^0 + x_4^0) \\ (2) & x_1^0 = P(B|A)(x_1^0 + x_2^0) \\ (3) & x_1^0 + x_2^0 + x_3^0 + x_4^0 = 1 \\ (4) & \forall i(x_i^0 \geq 0)\end{cases}$$

$$(5) \quad x_2^0 + x_4^0 = 0 \quad (1), (0.1)$$

$$(6) \quad x_1^0 + x_3^0 = 1 \quad (3), (4-6)$$

A solution for the system (S_0) is $P_0(x_1^0) = P_0(x_3^0) = .5$ and $P_0(x_2^0) = P_0(x_4^0) = 0$. Then, $P_0(B|\Omega) = 1$ is satisfied, since $P_0(B|\Omega) = \frac{x_1^0+x_3^0}{x_1^0+x_2^0+x_3^0+x_4^0} = \frac{.5+.5}{.5+0+.5+0} = 1$. Moreover, $P_0(B|A) = 1$ is satisfied, since $P_0(B|A) = \frac{x_1^1}{x_1^1+x_2^1} = \frac{.5}{.5+0} = 1$. Therefore, the assessment (0.1) and (0.2) on $\mathscr{C}$ is coherent. We note that in this proof equation (2) is irrelevant.

B.2 Negated Paradox 1

The NEGATED PARADOX 1,

From $P(B) = 1$ *infer* $P(\neg B|A) \in [0, 1]$

is used by Pfeifer and Kleiter (2011, see Section 3.3 on p. 59) to obtain a richer set of experimental tasks. The proof of the probabilistic non-informativeness consists of two parts. In part I we show that $P(\neg B|A)$ may be equal to zero if $P(B) = 1$, and in part II we show that $P(\neg B|A)$ may be equal to one if $P(B) = 1$.

Part I: Proof that $P(\neg B|A)$ may be equal to zero if $P(B) = 1$

Let Ω denote the certain event. Is the following assessment on the list of conditional events $\mathscr{C} = \{B|\Omega, \neg B|A\}$ coherent?

$$(0.1) \quad P(B|\Omega) = 1$$

$$(0.2) \quad P(\neg B|A) = 0$$

The set of the atoms $\mathscr{A}_0 = \{A_1, \ldots, A_4\}$ is generated by the following list of atoms A_i:

$$\begin{array}{rcrcr} A_1 & \Leftrightarrow & A & \wedge & B \\ A_2 & \Leftrightarrow & A & \wedge & \neg B \\ A_3 & \Leftrightarrow & \neg A & \wedge & B \\ A_4 & \Leftrightarrow & \neg A & \wedge & \neg B \end{array}$$

The system (S_0) is:

$$(S_0)\begin{cases} (1) & (x_1^0 + x_3^0) = P(B|\Omega)(x_1^0 + x_2^0 + x_3^0 + x_4^0) \\ (2) & x_2^0 = P(\neg B|A)(x_1^0 + x_2^0) \\ (3) & x_1^0 + x_2^0 + x_3^0 + x_4^0 = 1 \\ (4) & \forall i(x_i^0 \geq 0) \end{cases}$$

$$\begin{array}{lll} (5) & x_2^0 + x_4^0 = 0 & (1), (0.1) \\ (6) & x_2^0 = 0 & (2), (0.2) \\ (7) & x_1^0 + x_3^0 = 1 & (3), (4-6) \end{array}$$

A solution for the system (S_0) is $P_0(x_1^0) = P_0(x_3^0) = .5$ and $P_0(x_2^0) = P_0(x_4^0) = 0$. Then, $P_0(B|\Omega) = 1$ is satisfied, since $P_0(B|\Omega) = \frac{x_1^0+x_3^0}{x_1^0+x_2^0+x_3^0+x_4^0} = \frac{.5+.5}{.5+0+.5+0} = 1$. Moreover, $P_0(\neg B|A) = 0$ is satisfied, since $P_0(\neg B|A) = \frac{x_2^1}{x_1^1+x_2^1} = \frac{0}{.5+0} = 0$. Therefore, the assessment (0.1) and (0.2) on $\mathscr{C}$ is coherent.

Part II: Proof that $P(\neg B|A)$ may be equal to one if $P(B) = 1$

Is the following assessment on the list of conditional events $\mathscr{C} = \{B|\Omega, \neg B|A\}$ coherent?

$$\begin{array}{ll} (0.1) & P(B|\Omega) = 1 \\ (0.2) & P(\neg B|A) = 1 \end{array}$$

The set of the atoms $\mathscr{A}_0 = \{A_1, \ldots, A_4\}$ is generated by the following list of atoms A_i:

$$\begin{array}{lcrcr} A_1 & \Leftrightarrow & A & \wedge & B \\ A_2 & \Leftrightarrow & A & \wedge & \neg B \\ A_3 & \Leftrightarrow & \neg A & \wedge & B \\ A_4 & \Leftrightarrow & \neg A & \wedge & \neg B \end{array}$$

The system (S_0) is:

$$(S_0) \begin{cases} (1) & (x_1^0 + x_3^0) = P(B|\Omega)(x_1^0 + x_2^0 + x_3^0 + x_4^0) \\ (2) & x_2^0 = P(\neg B|A)(x_1^0 + x_2^0) \\ (3) & x_1^0 + x_2^0 + x_3^0 + x_4^0 = 1 \\ (4) & \forall i (x_i^0 \geq 0) \end{cases}$$

$$\begin{array}{lll} (5) & x_2^0 + x_4^0 = 0 & (1), (0.1) \\ (6) & x_1^0 = 0 & (2), (0.2) \\ (7) & x_3^0 = 1 & (3), (4-6) \end{array}$$

$P_0(B|\Omega) = 1$ is satisfied, since $P_0(B|\Omega) = \frac{x_1^0 + x_3^0}{x_1^0 + x_2^0 + x_3^0 + x_4^0} = \frac{0+1}{0+0+1+0} = 1$. Since x_3^0 is not necessarily equal to zero, we can delete equation (1) and construct the next system:

$$(S_1) \begin{cases} (1') & x_2^1 = P(\neg B|A)(x_1^1 + x_2^1) \\ (2') & x_1^1 + x_2^1 = 1 \\ (3') & \forall i (x_i^1 \geq 0) \end{cases}$$

$$\begin{array}{lll} (4') & x_1^1 = 0 & (1'), (0.1) \\ (5') & x_2^1 = 1 & (2'), (3', 4') \end{array}$$

$P_1(\neg B|A) = 1$ is satisfied, since $P_1(\neg B|A) = \frac{x_2^1}{x_1^1 + x_2^1} = \frac{1}{0+1} = 1$. Therefore, the assessment (0.1) and (0.2) on $\mathscr{C}$ is coherent. This concludes the proof that the NEGATED PARADOX 1 is probabilistically non-informative under the conditional event interpretation of the conditional.

We observe in part II of the probabilistic non-informativeness of the NEGATED PARADOX 1 a similar situation as in part I of the corresponding proof of PARADOX 1 (see p. 134f). The probability assessment (0.1) and (0.2) is coherent, but it violates the classical probability axioms. If $P(B) = 1$, then $P(\neg B|A) = \frac{P(A \wedge \neg B)}{P(A)} = \frac{0}{P(A)}$. $P(B|A) = 0$ if $P(A) > 0$, otherwise $P(B|A)$ is undefined.